Pedro Ravela | Beatriz Picaroni | Graciela Loureiro

¿CÓMO MEJORAR LA EVALUACIÓN EN EL AULA?

Reflexiones y propuestas de trabajo para docentes



Colección
APRENDIZAJES CLAVE
PARA LA EDUCACIÓN INTEGRAL

¿CÓMO MEJORAR LA EVALUACIÓN EN EL AULA?

Reflexiones y propuestas de trabajo para docentes

Pedro Ravela, Beatriz Picaroni y Graciela Loureiro

Colección
APRENDIZAJES CLAVE
PARA LA EDUCACIÓN INTEGRAL










¿CÓMO MEJORAR LA EVALUACIÓN EN EL AULA?

Primera edición, 2017

Argentina 28,
Centro 06020
Ciudad de México

Barranca del Muerto 341,
San José Insurgentes 03900
Ciudad de México

Editor
Fernando Díaz

Diseño
Patricia Carretto

ISBN de la colección: 978-607-97644-4-9
ISBN: 978-607-8558-39-1
ISBN 978-9974-8556-7-0 Grupo Magro Editores

Impreso en México
DISTRIBUCIÓN GRATUITA - PROHIBIDA SU VENTA

ÍNDICE

PREFACIO

Distinguidas maestras y maestros de educación básica

La Reforma Educativa en curso establece que todas las niñas, niños y adolescentes de México tienen derecho a una educación de calidad. Las medidas que se han diseñado para contribuir a este importante propósito se han incluido en la colección *Aprendizajes Clave para la Educación Integral*, con el objetivo principal de que los estudiantes aprendan y se desarrollen de manera integral para que vivan plenamente en la sociedad del siglo XXI.

Un concepto central de esta serie es el de *Aprendizaje Clave*, que se define como un conjunto de conocimientos, prácticas, habilidades, actitudes y valores fundamentales que contribuyen sustancialmente al crecimiento integral del estudiante y que se desarrollan específicamente en la escuela, los cuales, de no ser aprendidos, dejarían carencias difíciles de compensar en aspectos cruciales para su vida. Centrar el quehacer de las escuelas en los *Aprendizajes Clave* de los alumnos demanda ofrecer a los docentes recursos que les permitan evaluar con efectividad si cada estudiante ha adquirido o no dichos aprendizajes.

Por tanto, la SEP y el INEE consideran necesario acercar a los docentes información actualizada y de calidad que les permita evaluar mejor el desempeño de cada alumno, así como identificar las áreas de oportunidad para realizar intervenciones mejor informadas y que redunden en el éxito de todos los estudiantes. En ese sentido, *¿Cómo mejorar la evaluación en el aula? Reflexiones y propuestas de trabajo para docentes* aporta nuevos elementos de análisis y herramientas prácticas para transformar los procesos de evaluación que cotidianamente se realizan en las aulas.

Esta obra se incorporará al acervo de las bibliotecas de las escuelas de educación básica y normal, para que esté a disposición de todos los docentes. Se les invita a leerla y analizarla, y a que organicen círculos de estudio para obtener el mayor provecho de ella.

Secretaría de Educación Pública

Instituto Nacional para
la Evaluación de la Educación

ACERCA DE LOS AUTORES

Graciela Loureiro es profesora de educación primaria. Beatriz Picaroni también es docente, y además cuenta con un máster en Políticas Públicas y con un diplomado en Ciencias Sociales. Pedro Ravela es profesor de Filosofía de educación media y tiene un máster en Ciencias Sociales y Educación.

Los tres se han especializado en evaluación educativa a lo largo de los últimos 20 años. Colaboraron en los orígenes del sistema nacional de evaluación en lo que en aquel entonces (1995) fue la Unidad de Medición de Resultados Educativos de la Administración Nacional de Educación Pública (ANEP) de Uruguay. Durante varios años trabajaron en el desarrollo de lo que actualmente es la División de Investigación, Evaluación y Estadística (DIEE) de la ANEP, donde actualmente Graciela Loureiro se desempeña como coordinadora del Departamento de Evaluación de Aprendizajes. Pedro Ravela fue coordinador nacional del Programa Internacional para la Evaluación de Estudiantes (PISA) y Beatriz Picaroni fue directora de Investigación y Evaluación.

Los autores compartieron la docencia y la investigación en torno a la evaluación en el aula, en el marco del Instituto de Evaluación Educativa (IEE) de la Universidad Católica del Uruguay, del cual Pedro Ravela fue director entre 2007 y 2012. Juntos desarrollaron el Diploma en Evaluación de Aprendizajes en dicha institución, del cual Graciela Loureiro continúa siendo docente. Beatriz Picaroni y Pedro Ravela fueron además docentes en las maestrías en educación con énfasis en Currículo y en Evaluación de Sistemas Educativos, así como en el Diploma en Evaluación Docente. En forma individual y conjunta han desarrollado trabajos de investigación y consultoría en temas de evaluación educativa en varios países de la región y son autores de diversas publicaciones sobre el tema.

Pedro Ravela integró la Coordinación Técnica del SERCE en la Oficina Regional de Educación para América Latina y el Caribe de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (OREALC) en los años 2005 y 2006, y fue Director Ejecutivo del Instituto Nacional de Evaluación Educativa (INEEd) de Uruguay entre 2012 y 2014.

Actualmente Pedro Ravela se desempeña como docente e investigador independiente, Beatriz Picaroni es asesora del INEEd y Graciela Loureiro es coordinadora del Departamento de Evaluación de Aprendizajes en la ANEP.

AGRADECIMIENTOS

Queremos expresar en primer lugar nuestro agradecimiento a todos los docentes que, en el marco de los trabajos de investigación que realizamos en los años 2008 (en Lima, Montevideo, Buenos Aires, Cali, ciudad de Guatemala, Ciudad de México, San Salvador y San José de Costa Rica) y 2012 (en Cali, Lima, Montevideo y Santiago de Chile), nos permitieron acceder a sus clases, a sus materiales de trabajo y a sus preocupaciones en torno a la evaluación, y compartieron con nosotros sus reflexiones y sus propuestas para la evaluación de aprendizajes de sus estudiantes.

Del mismo modo, queremos expresar nuestro reconocimiento a todos los docentes y directivos con quienes tuvimos la oportunidad de interactuar y compartir reflexiones y experiencias en torno a la evaluación, durante los cinco años en que los autores trabajamos juntos en los Diplomas en Evaluación de Aprendizajes (2009 a 2013), realizados en Uruguay y Panamá, así como en talleres realizados en Aguascalientes, Buenos Aires, Cali, ciudad de Guatemala, Ciudad de México, ciudad de Panamá, Lima, Montevideo, Piura, Santo Domingo y San José de Costa Rica.

Agradecemos especialmente la colaboración de las siguientes instituciones con nuestros trabajos de campo: la Fundación Chile en Santiago; el Grupo de Investigación sobre Evaluación de la Calidad de la Educación (GIECE) en la Universidad Sanbuenaventura de Cali; la Unidad de Medición de la Calidad (UMC) de Perú; el Departamento de Educación de la Universidad Católica Argentina; el Instituto Nacional para la Evaluación de la Educación (INEE) de México; la Dirección de Gestión y Evaluación de la Calidad de Costa Rica; el Sistema Nacional de Evaluación de los Aprendizajes (SINEA) de El Salvador y la Dirección General de Evaluación Investigación y Estándares Educativos (DIGEDUCA) de Guatemala.

Queremos también hacer un reconocimiento explícito de los intercambios académicos que mantuvimos con el GIECE de la Universidad San Buenaventura de Cali, liderado por Dulfay Astrid González, y con el equipo de investigación de la Universidad Autónoma de Aguascalientes, dirigido por Felipe Martínez Rizo.

Reconocemos especialmente el apoyo brindado por Adriana Aristimuño, quien era decana de la Facultad de Ciencias Humanas de la Universidad Católica del Uruguay cuando se creó el Instituto de Evaluación Educativa y el Diploma en Evaluación de Aprendizajes.

Muchísimas gracias a la maestra Margarita Zorrilla, de larga trayectoria y reconocimiento en México, tanto en el campo pedagógico como en el de las políticas educativas, que aceptó revisar nuestro texto y escribir el prólogo.

Finalmente, queremos expresar un especial y entrañable agradecimiento, y reconocer, a Patricia Arregui por su apoyo y amistad, quien desde el Grupo de Análisis para el Desarrollo (GRADE) en Perú y desde su liderazgo en el Grupo de Trabajo sobre Estándares y Evaluación del Programa de Promoción de la Reforma Educativa en América Latina (PREAL), nos animó e impulsó a estudiar las evaluaciones en el aula y nos estimuló incansablemente en la realización de nuestras investigaciones y en la difusión de sus resultados en toda la región.

A Patricia y a todos los docentes les dedicamos este libro.

PRÓLOGO

Innovar desde la práctica para mejorar la evaluación en el aula[1]

Margarita Zorrilla Fierro[2]

Cuando los autores –Pedro, Beatriz y Graciela–, me pidieron hacer el prólogo de este libro no dudé ni un segundo en aceptar; no solo porque admiro profundamente su trabajo como docentes e investigadores, sino también porque la evaluación en aula me parece un asunto y un tema fundamental que requiere ser atendido de manera particular en la formación de docentes.

La evaluación educativa ha adquirido una importancia medular en los sistemas educativos de muchos países. Los de América Latina y el Caribe no son la excepción. Esto es así debido al enorme crecimiento de la cobertura de los servicios educativos. Hoy ingresan a la enseñanza primaria y media niñas, niños y adolescentes de familias que antes no iban a la escuela. La escolarización de la población ha sido un propósito y una tarea fundamental para nuestros países. Al ingresar a la escuela, aquellos que estuvieron excluidos de ella presentan una exigencia mayor para garantizar el ejercicio del derecho a ser educados. Hemos conseguido un éxito relativo al democratizar el acceso a la escuela, pero no hemos logrado de la misma manera la igualación en el acceso a los aprendizajes. De ahí el papel de la evaluación, pues permite iden-

1. Agradezco las valiosas aportaciones de Lídice Rocha Marenco y de Martha Elena Fuentes, colegas en el Instituto Nacional para la Evaluación de la Educación (INEE) de México, con quienes conversé ampliamente sobre el contenido de este libro y que me dieron mucho apoyo para escribir este pequeño prólogo a un gran libro.
2. Docente en todos los niveles educativos, desde la educación primaria hasta el doctorado. Investigadora de procesos, actores y resultados educacionales. Ha tenido responsabilidades en la gestión de la educación y de la evaluación a instituciones del Estado mexicano. Actualmente funge como consejera de la Junta de Gobierno del Instituto Nacional para la Evaluación de la Educación en su país.

tificar las brechas en muchos aspectos referidos a la realización del derecho de todos los individuos a recibir educación.

En las últimas décadas se extendió la evaluación externa a gran escala por medio través de programas y proyectos nacionales e internacionales. Sin embargo, hay una conciencia creciente de sus limitaciones o alcances, sobre todo por lo que esta evaluación signifique para los docentes y les sea de utilidad en su quehacer en las salas de clase. Por ello, en los últimos años ha empezado a cobrar mayor importancia la evaluación educativa dentro del aula, y vemos que la cantidad de estudios relacionados con el tema ha ido aumentando de manera significativa. Conocer lo que acontece en las aulas para incidir en su mejoramiento es un gran desafío para los docentes, académicos, investigadores y también para los políticos.

Si bien nuestra región ha dedicado grandes esfuerzos a estudiar y mejorar la evaluación de logro de aprendizajes a gran escala, poco se ha dicho sobre la evaluación en el aula, tal vez por su complejidad, pero lo cierto es que se trata de un tema que está pendiente en el debate educativo. Pedro, Beatriz y Graciela con este libro buscan aportar a este debate y, lo que es más, contribuir con soluciones.

Por lo antes dicho, celebro tener frente a nosotros un libro que no pretende ser un manual, la presentación de resultados de investigación, la narración de experiencias, ni un documento de orientación estrictamente teórica, sino más bien una especie de sincretismo entre estos elementos al entrelazarlos en una trama compleja por sí misma, pero sencilla en su tratamiento, que nos presenta una propuesta de proyecto para construir una nueva cultura de la evaluación en el aula. Pocos libros tienen un acercamiento tan propositivo. Por ello, es importante que seamos lectores activos, con apertura para reflexionar e incluso cuestionar sinceramente nuestras propias prácticas como docentes en las aulas o como formadores o como diseñadores del currículo o de materiales educativos. También el libro es una invitación a reflexionar sobre los problemas del sistema educativo en el que estemos inmersos, pero, sobre todo, a que tengamos la disponibilidad a introducir cambios, por pequeños que sean, para mejorar nuestro quehacer, sobre todo nuestras formas de enseñar.

El contenido de este libro trasciende las fronteras del país de origen de los autores, todos ellos uruguayos. Su pertinencia y actualidad conciernen a todos nuestros países de Iberoamérica, y yo diría que más allá. La sencillez con la que se presenta la complejidad de la evaluación en el aula en relación con la enseñanza y el aprendizaje es, sin duda, uno de los rasgos emblemáticos de esta obra.

Del Libro

Amable lector, sea usted docente, formador de docentes o tomador de decisiones en el ámbito de la formación de docentes, del currículo, de la evaluación educativa o simplemente un ciudadano interesado en la educación, tiene usted un gran libro entre sus manos. Es accesible a todo tipo de público, en gran medida por la sencillez del lenguaje y por los múltiples ejemplos pertinentes y entendibles sobre evaluaciones en el aula en varios países de la región, los cuales han sido recuperados y analizados por diversas investigaciones, debido a sus propuestas sobre cómo mejorar las evaluaciones, entre otras cosas.

Como lo he dicho, tiene ante usted un libro, antes que nada, propositivo. Desde las primeras páginas los autores nos invitan a realizar cambios y ajustes a nuestras formas de evaluar, pero sobre todo a pensar sobre la evaluación que usted o yo realizamos a nuestros alumnos. Tiene ante sí un ejercicio de análisis que posibilita la autocrítica al permitir identificar áreas de mejora como aliciente para un cambio fecundo.

Este libro que nos ofrecen Pedro, Beatriz y Graciela en ningún momento es prescriptivo. Sin embargo, logra guiar al docente con estrategias e ideas concretas hacia formas de evaluación más útiles para los docentes y los estudiantes, en tanto que ayuda a profundizar el aprendizaje significativo y busca atender la propuesta de enseñanza de manera más congruente. El lector se verá constantemente invitado a explorar diversas opciones de evaluación en lugar de conformarse con una sola, la que probablemente no le proporcione los resultados deseados ni cumpla con el propósito que busca como docente, esto es, que todos sus alumnos logren al máximo los aprendizajes que se espera de ellos.

Una de las propuestas más importantes del libro es, a mi juicio, la de fomentar el trabajo colectivo de reflexión y mejora continua para que a través del diálogo se vayan construyendo propuestas más sólidas que contribuyan a mejorar las prácticas de enseñanza y de evaluación de todos.

Aunque existen temas aparentemente tan cotidianos que parece que ya no se puede decir más al respecto o que no hay nada para debatir, los autores logran poner en duda la normalidad de la formulación de las actividades de evaluación, de la evaluación auténtica, de la subjetividad del proceso de calificación, y, por supuesto, de la evaluación formativa, la escurridiza del cuento.

Del sentido de la educación y la evaluación

Otro de los grandes atributos de este libro es que nos invita a preguntarnos ¿cuál es el sentido que le damos a la educación? Se trata de una pregunta que siempre es vigente y que a cada generación le corresponde resolver, ya sea a nivel mundial, regional, nacional, subnacional o de cada institución educativa, pues como sociedad hemos de precisar para qué educamos a nuestros niños y jóvenes. Por más que reformemos nuestros currículos, que mejoremos las capacidades de gestión de las escuelas y que llenemos de recursos tecnológicos nuestras aulas, si no tenemos claridad acerca de para qué educamos a nuestros niños y jóvenes, difícilmente conseguiremos mejorar de manera permanente la educación.

Partiendo de la idea de que la educación es un medio para conseguir un fin, no un fin en sí misma, resulta fundamental decidir cuál es el fin de que nuestros niños y jóvenes reciban una cantidad determinada de años de educación obligatoria. ¿Es para tener empleados más productivos?, ¿para reducir las brechas de desigualdad entre distintos sectores de población?, ¿para generar ciudadanos responsables?, ¿para qué?

Definir el sentido de la educación es la guía fundamental para determinar qué queremos que nuestros alumnos aprendan, cuáles son aquellos conocimientos clave, básicos para su desarrollo como personas, que consideramos que deben adquirir y desarrollar. Pero si no sabemos para qué, no podemos definir qué, haciendo inútil cualquier tipo de reforma curricular.

Esto nos trae a otro problema: si no sabemos para qué enseñamos, y por lo tanto no sabemos qué enseñar para lograr estos objetivos, ¿cómo vamos a elegir la mejor estrategia para enseñarlo?, ¿cómo vamos a evaluar que, en efecto, nuestros niños y jóvenes están aprendiendo lo que como sociedad consideramos fundamental?

Si partimos del supuesto de que se enseña en función del sentido y la finalidad que le logremos imprimir a la educación, esto nos lleva a otra complicación: la enseñanza depende de la interpretación y la apropiación de lo que cada docente entienda por el sentido de la educación y el propósito que esta debe cumplir, multiplicando así la cantidad de formas de enseñar posibles y de evaluar. Si bien la diversidad ofrece una gran riqueza, es necesario lograr consensos en cada sociedad sobre esos básicos irrenunciables del para qué educar.

Otro tema fundamental en esta obra es la oposición entre la evaluación como certificación/calificación y la evaluación como apoyo y fomento del aprendizaje. Los autores muestran que en la región los docentes de los distintos niveles educativos tienden a pasar más tiempo evaluando aprendizajes superficiales. Probablemente no porque no consideren importante el apren-

dizaje profundo, sino porque ellos mismos no saben cómo poner de relieve el nivel de aprendizaje y desempeño de sus estudiantes.

Esto me hace recordar el trabajo hecho por Benjamín Bloom conocido como la *taxonomía de los objetivos educacionales*. Su preocupación era conseguir "operacionalizar los objetivos de la educación", lo que le llevó a realizar una vasta y amplia investigación para lo que utilizó los exámenes que los docentes hacían y aplicaban a sus estudiantes para averiguar si habían aprendido o no. Este estudio le permitió construir seis categorías para agrupar las demandas de los docentes a sus alumnos, a saber: conocimiento, comprensión, aplicación, análisis, síntesis y evaluación.[3]

La educación no debería ser percibida por los estudiantes como conseguir buenas notas sin comprender el porqué de lo que se está aprendiendo y se lo olvide media hora después porque no se le encuentra el sentido o la utilidad. No podemos seguir dando más peso a las calificaciones que al aprendizaje profundo si queremos ciudadanos capaces de responder a los retos de la vida hoy día. En la actualidad, las sociedades democráticas y los espacios de trabajo requieren de individuos reflexivos y críticos que puedan poner en práctica sus conocimientos en escenarios nuevos y cambiantes, cosa que no estamos consiguiendo en este momento. Este no es un problema que solo se presente en nuestros países de la América Latina y en otras regiones en vías de desarrollo, es una situación más generalizada. Sin embargo, países desarrollados ya están tomando cartas en el asunto. Finlandia es un ejemplo.

Asimismo, concentrarnos en calificaciones no solo genera problemas de reprobación de curso o de tránsito de un grado a otro o de un nivel educativo a otro, sino que pueden tener efectos de largo plazo sobre las personas y sobre la calidad de la educación en el sistema escolar en su conjunto. Si el objetivo es que todos los alumnos "pasen de año", la calidad esperada de la enseñanza se relajará, o aquellos alumnos que no obtienen buenas calificaciones tendrán mayores incentivos para abandonar la escuela.

De la Evaluación

Los autores nos presentan el debate entre evaluación formativa y sumativa no como una división, sino como formas y propósitos de evaluar complementarios. A pesar de que en teoría todo el mundo parece estar a favor de una evaluación formativa, en la práctica se sigue dando mayor peso

3. Para ampliar la información recomiendo revisar al menos la introducción de la obra de Benjamín Bloom titulada *Taxonomía de los Objetivos Educacionales* publicada en su primera edición en 1956.

a la obtención de notas. En el discurso los docentes, académicos, evaluadores o investigadores pugnan por utilizar la evaluación formativa, sin embargo, en el terreno de los hechos, este discurso no se sostiene. Si bien es cierto que no se puede dejar de lado la evaluación sumativa, también es cierto que la finalidad de evaluar es la que debe definir el tipo de evaluación que se usa. En la actualidad existe una inercia en la forma de evaluar, no se cuestiona o se cuestiona poco, tanto al interior del aula como fuera de ella la forma en que deberían evaluarse los aprendizajes. Probablemente porque como docentes tenemos una intuición acerca de las implicaciones y requerimientos de una evaluación formativa, pero no sabemos bien cómo desarrollarla.

No hay una forma única, ni mejor para evaluar en el aula. Esta se debe construir a partir del contexto particular del docente y de sus estudiantes, de las necesidades que estos reflejen. Los mecanismos de evaluación no pueden ser fijos ni homogéneos porque cada grupo es diferente. La evaluación en aula debe adaptarse, como sea necesario, para reflejar el aprendizaje verdadero de nuestros alumnos, por lo que no existe tampoco una fórmula para crear evaluaciones auténticas. La mejor forma de evaluar es aquella que fomente el aprendizaje, la reflexión y la crítica de los estudiantes, así como del propio docente. No hay que limitarse a una sola, aunque es cierto que existen limitaciones administrativas y organizacionales que determinan el uso preponderante de evaluaciones con fines de calificación a lo largo de todo el ciclo escolar.

La evaluación, cualquiera que esta sea, tiene que enfrentar la complicación de definir sus criterios y explicitarlos y explicarlos a las personas que serán evaluadas. Esto casi nunca se lleva a cabo, pero sin duda es un paso necesario si queremos crear un proceso de enseñanza justo, fructífero, eficaz y eficiente. La especificación de los criterios reduce la incertidumbre de los estudiantes y elimina la posibilidad de generar malas interpretaciones respecto a lo que se espera de ellos y diluye la opacidad en el proceso de enseñanza. Además, esto les permite y los invita a que concentren sus esfuerzos en las áreas de aprendizaje que la evaluación considera importante, de ahí la necesidad de que la evaluación esté bien fundamentada con respecto a los objetivos de enseñanza, y estos, con el sentido de la educación que se haya definido y consensuado.

Aunado a esto, prácticas de transparencia en la forma de evaluar en el aula son una forma de incentivar la autoevaluación, la coevaluación y toda la práctica reflexiva que estas conllevan. Esto invita a los estudiantes a ser más proactivos, analíticos y reflexivos, dentro y fuera del aula, dotándolos además de instrumentos para la vida diaria que los ayudarán a enfrentar los retos del día a día.

Es de gran importancia que los estudiantes conozcan los criterios bajo los que serán evaluados, pero también es de suma importancia para el propio docente, quien a partir de estos puede verificar con mayor claridad que se estén cumpliendo los objetivos de aprendizaje, identificar las brechas entre lo que se quiere enseñar y lo aprendido, y ajustar sus actividades conforme sea necesario.

El docente debe tomar más una posición de entrenador que de jurado. Dar un dictamen no sirve de mucho, pero sí puede generar grandes mejoras en su desempeño el enseñar a los niños y jóvenes cómo pueden mejorar con base en sus errores. Los niños y jóvenes aprenden con ensayo y error; su incentivo no debe ser la nota, ni superar a sus compañeros, ni quedar bien con sus padres y docentes, sino aprender y en todo caso superarse a sí mismos a lo largo del tiempo. Al situarnos como jurados, los docentes estamos creando una distancia difícilmente superable entre nosotros y nuestros estudiantes. Sin duda la evaluación del aprendizaje con fines de calificación es necesaria y, probablemente, ineludible, pero no podemos dejar de lado la evaluación para el aprendizaje, esa que se usa todos los días de un modo similar al que utiliza un entrenador deportivo para ayudar a mejorar a los deportistas que tiene a su cargo.

El sistema de calificación, en particular, recibe por parte de los autores una fuerte crítica, o más bien, una llamada de atención a repensarlo. Aunque en el libro se reconoce su utilidad, e incluso necesidad al momento de certificar la obtención de ciertos conocimientos, también ponen a la luz, en repetidas ocasiones, la dificultad de reflejar el aprendizaje verdadero a través de las calificaciones. En ningún momento se pugna por dejarlo de lado, porque al final se trata de una limitante o condición del sistema educativo que difícilmente cambiará, pero sí se ponen de relieve sus incongruencias, efectos nocivos y sin sentidos, al tiempo que se invita constantemente al docente a utilizar más la evaluación formativa.

El libro hace una apología tanto de la evaluación formativa como de la evaluación auténtica. Más allá de los elogios que merecen estos enfoques, se trata de un tema que puede resultar de interés para el lector pues tiene aplicaciones mucho más fáciles de implementar en la evaluación del día a día. Así, la evaluación auténtica fomenta la motivación de los alumnos para seguir explorando y no perder la curiosidad necesaria para aprender a lo largo de la vida.

Este tipo de evaluación ayuda a verificar el aprendizaje significativo, entendido este como un aprendizaje profundo con el que los estudiantes pueden resolver problemas simulados de la vida real. Se trata de reproducir contextos reales para que los niños puedan aplicar lo aprendido, no lo me-

morizado. Es una forma de mostrarles la utilidad de los conocimientos y que ellos mismos se interesen por profundizarlos. Su utilización no implica desplazar la ejecución de ejercicios repetitivos puesto que tienen propósitos diferentes. Mientras que las repeticiones ayudan a fijar y automatizar conocimientos, la evaluación auténtica, debido a sus características formativas, ayuda a verificar el aprendizaje significativo y las áreas donde hace falta profundizar o reforzar. De hecho, su rasgo más distintivo es de origen formativo, sin embargo, también puede ser utilizada con fines de calificación o de certificación, bajo reserva de no exagerar en su uso.

De la Formación para la evaluación

Uno de los grandes desafíos o área de oportunidad para mejorar la evaluación en aula es la de incorporarla como un contenido prioritario en la formación inicial de docentes, quienes han de desarrollar herramientas teóricas, metodológicas y técnicas para llevar a cabo la evaluación de los aprendizajes de sus alumnos. También es una enorme área de oportunidad construir una oferta consistente e innovadora de formación continua de docentes en la que se incorpore la evaluación educativa como un componente fundamental de desarrollo profesional. Es conocido que la evaluación que implementamos los docentes es una reproducción de lo que vivimos como estudiantes, aunque no solo eso, sino que también cuentan las experiencias compartidas con otros docentes y la experiencia que se va adquiriendo a lo largo del tiempo.

El tema de la formación en materia de evaluación educativa aún no está explícitamente incluido en la agenda de nuestros sistemas educativos, aunque hay que destacar que existen iniciativas y experiencias que son dignas de ser conocidas y analizadas. Aunque en la actualidad se habla mucho acerca de la profesionalización docente alrededor de la formación, tanto inicial, como continua, la evaluación como asunto de la formación es un tema del cual se habla poco, pareciera que se asume que el docente sabe evaluar lo que aprenden sus alumnos y otros rasgos o atributos de su desarrollo y que la forma en que lo hace es la mejor.

Este vacío de formación para la evaluación en aula tiene como consecuencia problemas como estrés en alumnos y docentes, ambigüedad en los resultados esperados y subjetividad al momento de calificar, sobre todo en aquellos sistemas educativos que le otorgan una mayor importancia a las calificaciones.

Muchos de los docentes aprenden a construir pruebas y a calificarlas, generalmente con ayuda de sus compañeros maestros. Si bien esto es un hecho positivo, el problema es que se puede generar reduccionismo porque no

se sabe o no se exploran las posibilidades para evaluar, solo se replica lo visto y lo conocido, tanto en formato y contenidos como en centrar la evaluación en calificación y no en aprendizaje. Esta situación reduce las posibilidades de integrar innovaciones o mejoras en la práctica que reflejen el para qué educamos, qué enseñamos, cómo lo hacemos y qué logran aprender nuestros alumnos.

Cabe resaltar que, a pesar de que una parte del conocimiento del docente proviene de la interacción con sus pares, esto no significa que los docentes cuenten con los espacios y tiempos necesarios o recomendables para reflexionar de manera colectiva en torno al tema y aprovechar la riqueza de conocimiento que se genera cuando se dispone de estos espacios. Esto se debe, en parte, a la falta de tiempo pues cualquiera que haya sido docente sabe lo sobrecargada que está su agenda. Pero, también hay que decirlo, porque predomina una cultura individualista que no considera necesario compartir experiencias ni conocimientos, por el contrario, lo considera un riesgo. Un ejemplo que me gustaría resaltar son los colegiados de docentes que en cada centro escolar de cada país funcionan de diversas maneras y que suelen denominarse genéricamente como "consejos técnicos". Estos consejos son espacios que, en teoría, buscan generar la reflexión colectiva para solucionar los problemas individuales de los estudiantes y de la escuela como conjunto. Lamentablemente estos colegiados no siempre consiguen su cometido al convertirse en espacios para la atención de tareas administrativas. Una consecuencia de la falta de conversación pedagógica entre los docentes es la pérdida de conocer "mejores prácticas" tanto dentro como fuera de la escuela.

Para cerrar

En definitiva, se trata de un libro que busca romper con las inercias a través del diálogo y la reflexión. Dado que la enseñanza es una actividad cultural, es necesario redefinirla si queremos ver cambios en cómo es percibida tanto por los alumnos como por nosotros los docentes, de lo contrario seguiremos repitiendo los mismos patrones que no nos han permitido llegar a nuestras metas, siendo la más importante la realización del derecho de todos a recibir una educación de calidad.

El cuestionamiento, análisis y reflexión al que nos invitan los autores, es complicado porque los docentes tienen una carga enorme de trabajo y pocos incentivos para poner un esfuerzo adicional en su labor que, además, difícilmente les es reconocido. Tenemos que encontrar los incentivos dentro de nosotros mismos para llevarlos a cabo, por el bien de los niños y jóvenes que nos ha tocado formar. No obstante, es necesario hacer que estos temas sean

parte de la agenda pública para que los tomadores de decisiones del más alto nivel de los países también puedan ver las dificultades a las que se enfrenta el docente en el aula y se generen los incentivos y mecanismos necesarios, como es el caso de ampliar el desarrollo profesional con ofertas formativas pertinentes y consistentes.

Me gustaría terminar con una reflexión acerca de la propuesta de hacer comunidad al interior de la escuela. Sí, entre docentes, pero también de fomentarlo entre nuestros niños y jóvenes, para justamente ir modificando poco a poco la educación que acontece en cada aula, en cada escuela. Me parece que, más allá de los sistemas educativos, este es un tema central para la sociedad en general pues vivimos en una sociedad que más bien fomenta el individualismo, la discriminación y la eliminación de vínculos. Los educadores necesitamos impulsar una sociedad fundada en la equidad e inclusión, en la cohesión y la justicia.

Como sociedad, necesitamos trabajar juntos y enfrentar los retos de manera colectiva y colegiada. De lo contrario, continuaremos reproduciendo patrones de injusticia y desigualdad. Los autores, dentro de su ámbito de acción nos proponen justamente hacer esto y lo celebro enormemente. Hoy más que nunca tenemos que cuestionarnos acerca del papel de la educación y hacer a nuestros niños y jóvenes responsables de su propio aprendizaje para que se conviertan en adultos responsables de su sociedad. ¿Qué y cómo hacer hoy día para dar viabilidad a estos planteamientos?

Deseo cumplir con mi aspiración de conseguir que esta obra tenga muchos, muchos lectores en todos nuestros países de Iberoamérica, pero muy especialmente en nuestra América Latina.

Se hace honor a un libro cuando es leído y lo es aún más cuando de la lectura pasamos a la reflexión compartida con otros y luego transitamos a las acciones que comprendamos apoyan la transformación de nuestras prácticas en las aulas de cada centro escolar.

Ciudad de México, marzo de 2017

PRESENTACIÓN

Este libro ha sido escrito pensando principalmente en los docentes de la región, maestros y profesores de educación primaria y media, si bien puede resultar de interés también para docentes universitarios. Nuestra intención ha sido construir una herramienta de trabajo que ayude a los colegas docentes a reflexionar, revisar y mejorar sus formas de evaluar a los estudiantes. Pretende ser un aporte para el trabajo en el aula en los temas de evaluación. También ha sido pensado como una herramienta para quienes tienen a su cargo la formación de docentes, tanto la inicial como la formación en servicio. Querríamos, además, que el libro sirviese para sensibilizar a decisores de políticas educativas en torno a la importancia de la evaluación en el aula y a la necesidad de construir condiciones que hagan posible un trabajo docente más profesional, colaborativo e innovador.

Modificar las formas de "hacer las cosas" en el aula requiere una cierta dosis de humildad. La invitación al lector de este libro es a que ponga en cuestión por un momento la forma en que lleva adelante sus tareas en el aula, a partir de una mirada reflexiva sobre sus propias evaluaciones. Para ayudarle, nos proponemos abrir varias ventanas hacia aspectos de la evaluación en el aula que, a partir nuestra experiencia, de nuestras investigaciones y de nuestro intercambio con colegas docentes, creemos que podrían hacerse mejor. Seguramente, en función de la diversidad de trayectorias, formación y experiencia acumulada, cada lector encontrará aspectos distintos para revisar y mejorar.

Para la redacción de este libro empleamos cuatro fuentes principales. En primer lugar, nuestros trabajos de investigación sobre la evaluación en el aula, realizados en varios países de América Latina. En segundo lugar, nuestras lecturas sobre el tema. En tercer lugar, la experiencia de muchos docentes con los que hemos tenido la oportunidad de interactuar en el marco de cursos y talleres. Finalmente, nuestra propia experiencia como docentes, en la que nos enfrentamos, a la hora de evaluar, a los mismos dilemas e incertidumbres que el resto de los colegas.

En función de esta variedad de fuentes, este libro no es un manual de evaluación, ni un libro de teoría de la evaluación, ni un informe de investi-

gación, pero tiene algo de cada uno de ellos. A lo largo del mismo el lector encontrará reflexiones y análisis conceptuales sobre la evaluación, resultados de los trabajos de investigación realizados por los autores en torno a la evaluación en el aula en la región, aportes de trabajos sobre la enseñanza y la evaluación realizados en distintos países y, en cada capítulo, sugerencias y pistas de trabajo concretas, así como experiencias innovadoras realizadas por otros docentes.

El libro ha sido escrito en Uruguay (por lo que tiene referencias a la realidad educativa en nuestro país), pero pensando en los colegas de toda América Latina, cuyas experiencias hemos recogido y con quienes hemos tenido múltiples instancias de intercambio. Deliberadamente lo hemos escrito en primera persona del plural, porque si bien en este momento no ejercemos la docencia directa, lo hemos hecho durante casi toda nuestra carrera, en los niveles primario y medio al comienzo, y en la formación de docentes en el nivel terciario hasta hace pocos años. Por eso nos sentimos involucrados y compartimos las dificultades y perplejidades que todo docente tiene en el momento de evaluar.

Un libro refleja siempre la historia intelectual y personal de sus autores, sus inquietudes y preocupaciones. El contenido de este libro se explica en buena parte por una larga historia compartida entre los autores, trabajando juntos en los temas y problemas de la evaluación educativa.

Los azares de la vida hicieron que nos encontrásemos hace ya más de 20 años, a mediados de 1995, en la Unidad de Medición de Resultados Educativos, en Uruguay. Cada uno de nosotros venía de una experiencia diferente. No nos conocíamos con anterioridad. Los tres habíamos trabajado como docentes en el aula. En ese año nos tocó trabajar juntos (y con un equipo más amplio y de excelentes personas y profesionales), para organizar la primera evaluación censal de aprendizajes en 6º de educación primaria en Uruguay, que tuvo lugar a fines de 1996.

Nuestra preocupación central en aquel momento era que dicha evaluación fuese de alguna utilidad para los maestros y maestras. Veíamos como inconveniente y equivocado el modelo de pruebas orientadas a la publicación de rankings de escuelas que, desde Chile, en ese momento era una fuerte referencia en la región. Optamos por construir un modelo enfocado en la devolución de información a escuelas y docentes. Sus pilares principales fueron la confidencialidad de los resultados de cada escuela, la consideración del contexto sociocultural en el análisis de los mismos, el foco en pruebas que fuesen relevantes desde una perspectiva pedagógica, el análisis didáctico de los resultados y la realización, a partir de los mismos, de un programa de formación en servicio para los docentes de las escuelas de contextos desfavo-

recidos, que incluía una remuneración a los docentes por su trabajo adicional fuera del aula. En los años siguientes impulsamos otras experiencias de evaluación innovadoras.

Luego de trabajar durante muchos años en evaluaciones a gran escala -y sin dejar de hacerlo-, en el año 2008 tuvimos la oportunidad, gracias a la iniciativa e invitación de Patricia Arregui, que en ese momento lideraba el Grupo de Trabajo sobre Estándares y Evaluación, de llevar adelante una investigación sobre las evaluaciones que tienen lugar dentro de las aulas. Su preocupación era justamente que, mientras en nuestros países se invierte una gran cantidad de dinero, tiempo y energía en las evaluaciones a gran escala, nadie observa lo que ocurre con la evaluación dentro de las aulas.

Hicimos ese trabajo desde el Instituto de Evaluación Educativa creado en 2007 en la Facultad de Ciencias Humanas de la Universidad Católica del Uruguay. La investigación se enfocó en las concepciones y prácticas de evaluación en Lenguaje y Matemática en 6º de Primaria e incluyó escuelas y maestros de ocho países de la región. El propósito principal fue conocer cómo los docentes conciben y realizan la evaluación de sus estudiantes, a través de entrevistas en profundidad y del registro y estudio de las propuestas de evaluación que utilizan (pruebas, ejercicios, exámenes, talleres).

En el año 2012 tuvimos la oportunidad de realizar una segunda investigación, replicando la anterior, pero esta vez en 3º de secundaria básica y en materias vinculadas con las Ciencias Naturales (Biología, Física y Química). Esta segunda investigación, que contó con la participación de Julia Leymonié, Jennifer Viñas y Carmen Haretche, fue realizada en Cali, Lima, Montevideo y Santiago de Chile, en el marco del Fondo de Investigaciones Educativas (FIE) de PREAL.

Estas investigaciones nos hicieron re-conocer, volver a conocer, la importancia de la evaluación en el aula y el desafío que ella implica para los docentes. Nos hizo volver a centrar nuestras preocupaciones en el aula.

Las características de las investigaciones realizadas y algunos de sus principales hallazgos se presentan a lo largo del libro.

A partir del primer estudio (2008), se fueron sucediendo una serie de talleres y presentaciones de los resultados en las Ciudades de Guatemala, Panamá y México, en Lima, Piura, Cali, Aguascalientes, Montevideo, Buenos Aires, Santo Domingo y San José de Costa Rica. En todas estas instancias tuvimos la oportunidad de profundizar en el intercambio con docentes y en nuestra comprensión de la problemática y las prácticas de evaluación en la región. A partir de la presentación y discusión de nuestros principales hallazgos, los fuimos revisando y re-significando, gracias a las perspectivas de los diferentes actores.

Como resultado casi natural de la experiencia antes relatada, en el año 2009 creamos el Diploma en Evaluación de Aprendizajes, que aún se continúa ofreciendo en la UCU. Se trata de una propuesta de posgrado para docentes en actividad, de carácter semipresencial y de un año de duración. En el marco de la misma los docentes participantes tienen la oportunidad de revisar sus prácticas y herramientas de evaluación en forma colaborativa con los docentes del curso y con otros colegas. Durante los años 2011 y 2012 se ofreció una edición especial de este Diploma en Panamá, a través del Servicio Nacional de Ciencia y Tecnología (SENACYT) de ese país. Estas experiencias fueron sumamente enriquecedoras para nosotros, por la oportunidad que significaron para intercambiar miradas y experiencias sobre la evaluación con una gran cantidad de docentes de todos los niveles de enseñanza.

Este libro es fruto de esta historia. A lo largo del mismo presentamos en forma organizada elementos y reflexiones conceptuales sobre la evaluación, evidencia empírica sobre los enfoques y prácticas de evaluación de los docentes de la región y propuestas concretas para trabajar en equipo en la revisión y mejora de dichas prácticas.

En el capítulo 1 hacemos una primera aproximación conceptual al tema de la evaluación en el aula y a las dificultades para modificar las prácticas de enseñanza. Describimos, además, los trabajos de investigación que sustentan el libro, cuyos principales resultados se incluyen luego en los diferentes capítulos.

El capítulo 2 está dedicado al análisis de los procesos cognitivos involucrados en las tareas de evaluación. La pregunta central para el lector en este capítulo es: *¿qué requiere de mis estudiantes responder a las pruebas y a otras actividades de evaluación que les propongo?* A lo largo del mismo se desarrolla una herramienta para analizar los componentes centrales de una actividad de evaluación y los diversos procesos de pensamiento que están involucrados en la resolución de las mismas. Se analizan las diferencias entre aprendizaje superficial y aprendizaje en profundidad, así como diferentes caminos para avanzar de una enseñanza orientada a la memorización y reproducción de contenidos y procedimientos, hacia una enseñanza orientada a la comprensión y la reflexión.

El capítulo 3 desarrolla una propuesta de evaluación y enseñanza a través de situaciones auténticas. Estas implican colocar los contenidos que enseñamos en el marco de un contexto que se asemeje a la forma en que el conocimiento es construido y utilizado en las diferentes disciplinas y en la vida social. La pregunta central para el lector en este capítulo es: *¿en qué medida presento a mis estudiantes los contenidos del programa de una manera que los vincule con las situaciones reales en las que dichos contenidos son relevantes*

y con las formas de producción de conocimiento propias de la disciplina en cuestión? En este capítulo el lector encontrará pistas de trabajo para transformar las pruebas y escritos que se utilizan habitualmente, en situaciones más ricas, motivadoras y desafiantes para los estudiantes. A lo largo del capítulo se analiza el concepto de evaluación auténtica y se muestran experiencias desarrolladas para el aula por docentes de educación primaria y media.

El capítulo 4 se enfoca en las prácticas de evaluación formativa. Este tema involucra una de las principales contradicciones en que incurrimos la gran mayoría de los docentes: enfatizamos la importancia de la evaluación formativa pero en la práctica pasamos la mayor parte del tiempo dando calificaciones o señalando errores en los trabajos de nuestros estudiantes. La pregunta central para el lector en este capítulo es: *¿qué tan incorporadas están en mi trabajo cotidiano en el aula las prácticas orientadas a realizar devoluciones a los estudiantes que les permitan reflexionar sobre lo que están aprendiendo?* El capítulo enfatiza la necesidad de diferenciar las evaluaciones que tienen por finalidad calificar a los estudiantes, de las instancias y acciones de evaluación orientadas a apoyarlos en el proceso de aprendizaje. Se trabaja especialmente el conceptos de "devolución" y se ofrecen pistas de trabajo para incorporar la evaluación formativa al trabajo cotidiano, analizando las maneras de formular preguntas en clase y los modos de corregir y devolver trabajos y pruebas. Se ofrecen además sugerencias para promover de manera efectiva instancias de auto-evaluación y co-evaluación.

El capítulo 5 está destinado a analizar los criterios y procedimientos a través de los cuales calificamos a nuestros estudiantes. Dos son las preguntas centrales para el lector en este capítulo: i) *¿qué tan claros son para mis estudiantes los criterios y procedimientos que utilizo para calificarlos?*; y, ii) *¿qué evidencias reciben sobre lo que aprendieron y saben hacer a partir de las calificaciones que les damos?* A lo largo de este capítulo se analizan los reglamentos de evaluación vigentes en varios países de la región y el modo en que estos determinan un conjunto de prácticas de calificación cuya consecuencia principal es generar una cultura en la que los estudiantes están más orientados a las calificaciones que al aprendizaje. El capítulo analiza cómo aquellas -y su consecuencia directa, la reprobación de cursos- son un elemento constitutivo del modelo escolar propio de la sociedad industrial, a partir de lo cual se plantea la necesidad de pensar en un modelo alternativo. El capítulo se cierra con un conjunto de sugerencias para mejorar la forma en que manejamos las calificaciones en nuestros cursos.

Los capítulos 2 al 5 se cierran con un conjunto de pistas para el trabajo en los centros educativos, en torno a la mejora de las prácticas de evaluación relacionadas con el tema foco de cada capítulo. En los capítulos 4 y 5 se pre-

sentan las "rúbricas", en tanto herramientas útiles para la evaluación formativa (en el capítulo 4) y para la definición y comunicación de calificaciones (en el capítulo 5).

Solo nos queda invitar a los lectores a trabajar con otros colegas para continuar mejorando sus prácticas y propuestas de evaluación. Nuestra intención es realizar un aporte que, ojalá, ayude a cada lector a pensar en sus evaluaciones y en cómo podría mejorarlas. No tenemos ninguna pretensión de establecer un enfoque "fundacional" sobre la evaluación, ni proponer que habría que cambiar todo para evaluar bien, ni establecer una nueva "moda". Creemos que no hay un único y mejor modo de evaluar. Hay variedad de enfoques, que siempre son mejorables. Cada docente va construyendo su modo de actuar en este terreno a lo largo de su carrera. Para algunos, las propuestas incluidas en este libro servirán para dar un paso más en un camino ya iniciado. Para otros implicará revisar en profundidad rutinas y prácticas arraigadas a lo largo de mucho tiempo. La única sugerencia que podemos dar es animar a probar cambios concretos, apoyándose en algunas de las pistas de trabajo propuestas a lo largo del libro, sin pretender modificar de un día para otro lo que se viene haciendo. Como dice un antiguo refrán chino: "un camino de mil kilómetros comienza con un paso".

Esperamos que el texto contribuya a motivar a los lectores para ensayar nuevas formas de evaluar y de enseñar. Para algunos el desafío será más grande, para otros se tratará simplemente de avanzar en un camino en el que ya están trabajando. Si cada lector intenta introducir alguna innovación en sus clases, escribir este libro habrá valido la pena.

CAPÍTULO 1
EL ELEFANTE INVISIBLE

Evaluar es una tarea ingrata. La mayoría de los docentes preferiríamos sencillamente enseñar y no tener que evaluar. Los sentimientos negativos relacionados con la evaluación son variados. Entre los principales están, en primer término, el vinculado a la sobrecarga de trabajo y el cansancio que implica. Evaluar requiere revisar y corregir grandes cantidades de tareas, normalmente en horarios en los que, se supone, no estamos trabajando (noches, feriados y fines de semana). Siempre tenemos "tarea pendiente": cuando terminamos de corregir una tanda de trabajos, ya tenemos otra por delante. Un segundo sentimiento es el de aburrimiento, derivado del hecho de corregir varias decenas de trabajos similares (varios centenares, en la educación media). Leer repetidamente las mismas respuestas o corregir los mismos ejercicios es algo tedioso. Un tercer tipo de sentimiento que muchas veces está presente, es el de frustración. Esta se deriva del hecho de que al evaluar solemos constatar la distancia que existe entre lo que intentamos enseñar y el desempeño de nuestros estudiantes. En cuarto lugar cabe mencionar el sentimiento de incertidumbre y aún de cierto estrés o angustia. Esto suele ocurrir en las instancias de evaluación en las que tenemos que tomar decisiones vinculadas a la aprobación o reprobación de un curso. En esos momentos nos preguntamos si estamos evaluando en forma apropiada y si las decisiones que tomamos, que afectarán la vida de nuestros estudiantes, son justas y tienen el sustento adecuado. La pregunta "¿lo estaré haciendo bien?" sobrevuela nuestras evaluaciones en forma más o menos permanente.

Muchos de nosotros preferiríamos no tener que evaluar y no lo haríamos, si no fuese un requerimiento administrativo y una forma de "motivar" -¿obligar?- a los estudiantes a estudiar y a realizar las tareas propias del aprendizaje. Vivimos en una cierta contradicción. No nos gusta evaluar y no sabemos del todo cómo hacerlo bien, pero calificamos o asignamos puntajes la mayor parte del tiempo. Sostenemos en el discurso la importancia de la evaluación formativa por sobre las calificaciones, pero continuamente ponemos y registramos "notas". Esto ocurre por varias razones. Por un lado, porque es un requisito administrativo. Por otra parte, porque es una demanda constante de los propios estudiantes y de muchas familias. Y, además, porque en

cierta medida percibimos que es la única forma de lograr que los estudiantes se involucren con las tareas que les proponemos.

Por lo general no hemos tenido la oportunidad de formarnos adecuadamente para evaluar. La evaluación suele estar casi ausente, tanto durante las carreras de formación inicial como en los programas de formación en servicio. Cuando hay alguna instancia de formación, esta suele ser poco relevante para la práctica. Se enseñan algunos conceptos básicos (tales como la distinción entre evaluación formativa, sumativa y diagnóstica), algunos discursos teóricos sobre la misma y algunos conceptos de estadística que no suelen tener aplicación en la vida real en el aula (tales como la media, el modo, la mediana, la distribución de frecuencias y curva normal). En esta situación, cada docente "aprende" a evaluar como puede, en forma más o menos intuitiva, en cierta medida reproduciendo las formas en que fue evaluado cuando era estudiante y, a veces, recurriendo a materiales prestados de algún colega.

Con el paso de los años y la acumulación de experiencia, vamos desarrollando nuestra capacidad de valorar la diversidad de desempeños de nuestros estudiantes y cierta dosis de "buen ojo" para evaluar. Mientras algunos docentes se apoyan más en la asignación de puntajes a distintas tareas y ejercicios, otros optan por una aproximación más basada en sus percepciones del desempeño de los estudiantes que en puntajes matemáticos. Pero a todos nos cuesta explicitar a los estudiantes con claridad qué esperamos de ellos y qué criterios emplearemos para evaluarlos. El modo en que definimos las calificaciones suele ser opaco y difícil de comprender para los estudiantes, por más que interiormente nosotros tengamos cierto grado de certeza con respecto a la calificación que asignamos. También tenemos ciertas dudas existenciales con respecto a qué cosas valorar y qué peso darle a cada una: ¿debemos tener en cuenta el trabajo realizado por el estudiante o valorar el esfuerzo?; ¿hasta qué punto mantener la exigencia alta?; ¿en qué circunstancias o ante qué estudiantes debemos ser más "tolerantes"?

Este libro intentará clarificar algunos aspectos clave de la evaluación de aprendizajes en el aula y ofrecer pistas de trabajo que permitan abordar dicha tarea de un modo más satisfactorio para el docente y más justo con los estudiantes, y que sea consistente con nuestras intenciones educativas y con los aprendizajes que querríamos que desarrollen nuestros estudiantes.

1. El elefante invisible

En un reciente libro en el que realiza una completa e interesante revisión de las políticas educativas desarrolladas en América Latina durante la última década y media, Axel Rivas (2015) destaca que *"los exámenes y las*

calificaciones son el esqueleto de los sistemas educativos" y los describe como un "elefante invisible", haciendo alusión a la vez a la importancia que tienen y a la poca atención que se les presta. De Axel Rivas tomamos esta expresión como título para este capítulo.

En un apartado dedicado a analizar las políticas de evaluación de los estudiantes, Rivas se pregunta qué ocurrió en la región con los sistemas de promoción de los alumnos y las pruebas, reglamentos y requisitos para aprobar los cursos desde el comienzo del nuevo siglo. Su conclusión es que este tema ha recibido escasa atención y casi no ha sido estudiado, a pesar de su importancia crucial en la determinación de los aprendizajes y de la experiencia educativa de niños y jóvenes.

"El esqueleto del sistema educativo es un elefante invisible. Pocos estudios de política educativa hablan de los regímenes académicos de la educación primaria y secundaria. Es un tema escasamente investigado en relación con su peso en la vida de los sistemas y, sobre todo, de los alumnos. Basta ver qué preguntan la mayoría de ellos cuando comienza un curso en cualquier tipo de institución educativa del mundo: ¿cómo seremos evaluados?" (Rivas, 2015: 101).

Los sistemas de calificaciones juegan un papel central en la educación formal tal como está organizada en la mayor parte de nuestros países. Son un factor determinante de la experiencia educativa de los estudiantes y de su concepción acerca de en qué consiste su "oficio" como tales, es decir, aprender. Lo que evaluamos a través de pruebas y exámenes constituye un fuerte mensaje que comunica a estudiantes y familias qué esperamos que sean capaces de hacer y qué es valioso para nosotros -los docentes- como logro educativo. Por otra parte, mientras los docentes tendemos, al menos en el discurso, a dar más importancia a la función formativa de la evaluación que a las calificaciones, para los estudiantes evaluación es, ante todo, sinónimo de calificaciones o notas. Estas se han constituido en una pieza clave de los sistemas educativos que determina fuertemente toda la experiencia escolar, a pesar de que, según analizaremos en el capítulo 5, constituyen una rémora del modelo escolar propio de la sociedad industrial.

Según muestra Rivas en su estudio, una de las principales contradicciones que se han puesto de manifiesto en las políticas educativas en los últimos quince años, es la derivada de la tensión entre garantizar aprendizajes comunes a todos y atender a la diversidad de los estudiantes de un modo que les permita avanzar en sus trayectorias educativas. Esta tensión no siempre ha sido bien resuelta. Las principales medidas de política educativa en los países de la región en estos temas, han sido la eliminación de la repetición en algunos grados de primaria y la flexibilización de los regímenes de pasaje de grado en la educación media.

"Quizás, el mayor cambio fue la emergencia de nuevas instancias de evaluación y aprendizaje, conocidas como períodos de compensación o recuperación. *Se instalaron las recuperaciones pedagógicas, como períodos breves antes del inicio del ciclo escolar, donde los alumnos tienen clases antes de ser evaluados, o las evaluaciones preventivas durante el curso escolar, para detectar situaciones problemáticas y atenderlas dentro de la cursada con adecuaciones de pedagogía diferenciada. Todo esto trajo diversas respuestas docentes, políticas y sociales. Muchos consideraron que las nuevas posibilidades empujaban a los docentes a aprobar a los alumnos a cualquier costo. Eran los que tenían menos vocación de revisar las pedagogías o educar en la diversidad. La creencia en la meritocracia como criterio llevaba a la repetición y a la expulsión como medidas necesarias para mantener el orden en las aulas. Sin ellas, todo el reino escolar se pervertía en el caos del facilismo. Fue arduo el trabajo de muchos pedagogos, funcionarios y colectivos docentes, incluidos la mayoría de los sindicatos docentes de la región para evitar una confusión decisiva: que todo lo que implica dar nuevas posibilidades a los alumnos fuese visto como sinónimo de facilismo. En realidad, la propuesta era exactamente lo contrario: un trabajo arduo para enseñar más, revisar las pedagogías, personalizar la enseñanza y generar interés en los alumnos por su propia escolarización. Pero, muchas veces, ese mensaje llegaba de manos de gobiernos lejanos a las vivencias diarias de las aulas. Frías normativas que poco podían hacer con décadas de naturalización pedagógica de la exclusión. En muchas escuelas, la falta de recursos o de nuevas prácticas de formación docente para asumir estos desafíos pedagógicos se manifestaban de lleno en los intentos de hacer política por vía de la normativa. En muchos casos, esto generó frustración y resistencia, en uno de los capítulos claves de las brechas entre las políticas y las prácticas pedagógicas"* (Rivas, 2015: 105).

Desde nuestra perspectiva el análisis de Rivas peca de cierto "exceso" de optimismo. En primer lugar, porque en cierto modo plantea que la tensión ha sido resuelta cuando en general continúa instalada, de la misma manera que, con o sin razón, persisten las resistencias a la flexibilización de los regímenes académicos. En segundo lugar, porque entendemos que "la revisión de las pedagogías" (tal vez más bien de las didácticas), la "personalización de la enseñanza" y la generación de "interés en los alumnos por su propia escolarización" son aun tareas pendientes. Como veremos a lo largo de este libro, si bien existe la preocupación por introducir diversidad en las formas de enseñar y por asegurar las trayectorias de los estudiantes, estamos lejos aún de lograr cambios sustantivos en las formas de trabajar dentro del aula, en nuestras concepciones de qué son saberes relevantes y en los modelos de organización de las instituciones educativas.

Estos tres elementos van de la mano: la concepción de qué merece la pena ser aprendido y cómo; las formas de evaluar a los estudiantes y sus aprendizajes; y los modos de organizar la institución escolar en lo relativo a la estructuración de los tiempos cotidianos, los espacios físicos, las formas de agrupar a los estudiantes, los períodos anuales de cursos, las relaciones entre estudiantes y docentes, y los momentos en los que se determina si un estudiante ha logrado o no los aprendizajes esperados (así como la propia definición de "aprendizaje esperado") y su consecuencia inmediata, la aprobación o reprobación de un curso.

En este sentido, el elefante invisible sigue siéndolo y es más grande que lo planteado por el autor. Mientras el trabajo de Rivas utiliza esta imagen para referirse específicamente en los regímenes de aprobación y reprobación de cursos, en este libro analizaremos más ampliamente las prácticas de evaluación en el aula, tanto las que tienen por finalidad establecer una calificación como aquellas que tienen una finalidad formativa y que están estrechamente imbricadas con la propuesta de enseñanza del docente. Ambas reflejan nuestra forma de enseñar y lo que esperamos de nuestros estudiantes, ambas moldean la percepción y las actitudes de los estudiantes hacia la educación.

La evaluación en el aula constituye uno de los pilares de la cultura escolar dominante. Y, sin embargo, en general la ignoramos. No suele ser objeto de investigación y tiene un lugar absolutamente secundario en la formación de docentes, tanto en la inicial como en los programas de formación en servicio y desarrollo profesional. Es un elefante invisible. Mientras se invierten grandes cantidades de tiempo y dinero en la realización de evaluaciones a gran escala, son casi inexistentes los recursos y la atención destinada a mejorar las evaluaciones que ocurren cotidianamente dentro del aula. Pero el impacto de estas últimas sobre la enseñanza y sobre el aprendizaje es enorme, seguramente mucho mayor que el de las evaluaciones externas. Y nosotros, los docentes, seguimos haciendo las cosas como siempre se hicieron, asumiendo como "naturales" las formas establecidas de evaluar a los estudiantes.

El propósito de este libro es poner sobre la mesa las prácticas de evaluación, en tanto emergente o indicador que muestra qué y cómo enseñamos. Creemos que colocar nuestras propuestas y prácticas de evaluación como objeto de análisis y discusión es un excelente punto de partida para iniciar un proceso colectivo de revisión y cambio en las mismas.

2. La evaluación como forma de conocimiento

Evaluar es una actividad esencial y natural para el ser humano. Continuamente estamos evaluando para tomar decisiones de distinto tipo, tanto

a nivel individual como colectivo. Normalmente toda decisión de cierta importancia, así como buena parte de las decisiones triviales, están precedidas de alguna forma de evaluación.

A nivel individual o familiar, decisiones como la elección de un servicio de salud, el alquiler de una vivienda o la compra de una prenda de vestir, están precedidas por un proceso en el cual identificamos las distintas alternativas existentes y buscamos información sobre cada una de ellas: los servicios que se ofrecen; los costos de distintas alternativas; en el caso del servicio de salud, en qué lugares físicos hay servicios disponibles puede ser un dato importante, porque implica traslados; la distancia al lugar de trabajo puede ser un elemento importante para elegir vivienda, así como la cantidad de dormitorios y la iluminación; en el caso de la adquisición de una prenda de vestir, observaremos el diseño, los colores y, seguramente, nos la probaremos y observaremos "cómo nos queda". A medida que recogemos este tipo de información sobre las distintas alternativas, la iremos combinando con ciertas valoraciones: qué cosas priorizamos en un servicio de salud, según cuál sea nuestro estado físico; qué tipo de vestimenta nos resulta más agradable; qué ubicación, tamaño y estilo de vivienda queremos para vivir; cuánto estamos dispuestos a pagar; por mencionar algunas. Como resultado de esta interrelación entre información y valoraciones, llegaremos en cada caso a un juicio de valor o conclusión valorativa -*"esta es la mejor opción dentro de mis posibilidades", "este servicio es mejor que el resto teniendo en cuenta el costo"*-, que nos permitirá tomar una decisión.

A nivel social la evaluación está presente, de manera explícita o implícita en todas las decisiones colectivas y políticas. Evaluamos antes de decidir invertir en la ampliación de las actividades de una empresa, para resolver a cuánto debe incrementarse el salario mínimo nacional o antes de introducir cambios que permitan mejorar la implementación de un programa para erradicar el trabajo infantil. También en estos casos las decisiones están precedidas por la recogida y análisis de información y datos (en mayor cantidad y en forma más sistemática que en los ejemplos de la vida personal), que se combinan con valoraciones éticas y políticas, para llegar a juicios de valor del tipo: *"es mejor no ampliar las actividades de la empresa en este momento porque es muy riesgoso"; "es deseable incrementar el salario mínimo nacional hasta tal cifra pero hacerlo por encima de la misma será contraproducente para el conjunto de la economía"; "sería necesario mejorar el programa de erradicación del trabajo infantil introduciendo cinco cambios principales en su forma de operar"*.

Si el lector analiza con cuidado cada uno de los casos propuestos, notará que en todos ellos los juicios de valor resultantes y las decisiones a las que conducen son "discutibles", es decir, no son las únicas posibles. Utilizando

la misma información de base, dos personas distintas pueden llegar a juicios de valor y a decisiones diferentes sobre el servicio de salud a contratar, sobre la vivienda a alquilar, o sobre la prenda de vestir a comprar. De la misma manera, diferentes especialistas o políticos pueden llegar a juicios de valor y decisiones diferentes acerca de cuál podría ser la cifra adecuada para el salario mínimo nacional o cuáles deberían ser los principales cambios a introducir en el programa de erradicación del trabajo infantil. En el caso de la empresa, distintas personas dentro de la misma podrían considerar que el riesgo de la inversión está dentro de lo razonable y que este sería un buen momento para realizarla.

Esta es una característica central de los procesos de evaluación: sus "conclusiones" no son únicas ni indiscutibles, dado que se apoyan en información y en valoraciones. Como consecuencia de ello, a partir de los mismos datos, las "conclusiones evaluativas" pueden variar en función de los valores que se toman en consideración.

Este hecho suele dar lugar a una confusión: considerar que la evaluación es algo meramente subjetivo, una cuestión de gustos o preferencias, en cierto modo aleatorio o antojadizo. No lo es, en la medida en que tampoco lo son los valores éticos o estéticos. Los valores pueden ser diversos, pero no por ello son algo meramente caprichoso. Tampoco lo son las valoraciones que emergen de un proceso de evaluación: si la evaluación ha sido realizada de un modo apropiado, sus conclusiones serán consistentes con la evidencia empírica empleada y con los referentes valorativos considerados, si bien variarán en función de estos últimos.

Michael Scriven[1] (2013) define a la evaluación como "*el acto o proceso cognitivo por el cual establecemos una afirmación acerca de la calidad, valor o importancia de cierta entidad*". Dicha entidad, a la que denomina "evaluando", puede ser un objeto, un programa, un curso de acción, un desempeño, entre otros. Según Scriven, es necesario combatir la idea de que los valores son esencialmente subjetivos, una cuestión de gustos idiosincráticos, no contrastables o esencialmente imprecisos y cualitativos (32). El hecho de que "*el contenido de las afirmaciones evaluativas es extremadamente dependiente del contexto, difiriendo de un usuario a otro o de una situación a otra para un mismo usuario, no significa que su significado no sea claro*" (25).

1. Michael Scriven ha dedicado su vida a la evaluación. Nacido en Inglaterra en 1928, formado en Australia y radicado en los Estados Unidos hace varias décadas, es hoy un hombre cercano a los noventa años de edad que continúa realizando contribuciones importantes a esta disciplina, que él denomina "transdisciplina", porque atraviesa a todas las disciplinas, tanto científicas como artísticas.

De acuerdo con Scriven (2011), la evaluación es una forma de conocimiento, decisiva para la especie humana. *"Los seres humanos primitivos fueron evaluadores prácticos de todo lo existente"* (17). Probablemente algunas de las primeras inferencias evaluativas estuvieron vinculadas a los frutos a ingerir como alimento y a la selección de piedras apropiadas para la construcción de instrumentos. La elección y el descarte paralelo de piedras para construir instrumentos *"implican el empleo de estándares de aceptabilidad"* (17). A partir de la observación de distintas características de las piedras, los humanos debieron llegar a un juicio de valor, caso por caso, de cuáles podrían servir y cuáles no, o cuáles serían mejores que otras.

Desde ese hipotético comienzo, la evaluación ha sido una forma de conocimiento determinante y permanente en la adaptación y evolución de los seres humanos a lo largo de su historia sobre el planeta. *"La habilidad para evaluar es una parte importante del cuerpo de conocimiento que hemos aprendido fuera de la escuela. Se aprende por ensayo y error e inferencia y, luego de mucha experiencia (que puede ser supervisada), se internaliza como una habilidad perceptiva"* (27). Además de su uso cotidiano de un modo no sistemático por todas las personas, se transformó en una disciplina y en una profesión, que tiene sus procedimientos propios y sus requerimientos técnicos y éticos. *"La evaluación es una de las formas principales de ciencia social aplicada"* (21). En realidad, afirma el autor, la evaluación es una transdisciplina (11, 36), es decir, una disciplina que aporta herramientas a las restantes (36) (otros ejemplos de transdisciplina son la estadística y el diseño). En este sentido, la evaluación está presente al interior de todas las disciplinas y áreas de la actividad humana: en la producción de conocimiento en las diferentes ciencias, en las artes, en los deportes, en la educación, en la medicina, en la ingeniería, en la economía y en la política, por mencionar las principales. En todas ellas es necesario, en forma continua, tomar decisiones basadas en evidencia y valores, lo cual constituye la esencia de la evaluación.

En las distintas disciplinas y ámbitos de actividad humana la evaluación tiene siempre una misma lógica básica (Ravela 2003 y 2006; Scriven 2013):

- el punto de partida es siempre la existencia de ciertos propósitos o la necesidad de tomar una decisión o curso de acción;
- para ello es necesario establecer una afirmación valorativa sobre la situación, sobre los cursos de acción posibles o sobre la realidad evaluada;
- para ello se toman en consideración ciertos aspectos, referentes o "criterios de valor" considerados relevantes (que denominaremos dimensiones); cada dimensión tiene una importancia relativa diferente de las otras (algunas son más importantes que otras);

- recogemos alguna forma de evidencia empírica sobre cada una de las dimensiones definidas como relevantes con el fin de estimar el valor de cada una de ellas;
- la "conclusión evaluativa" o juicio de valor es el resultado de "coligar" las dimensiones, la evidencia y las estimaciones de valor, teniendo en cuenta la importancia relativa de cada dimensión.

El término "coligar" ha sido acuñado por Scriven (2013: 25-28). Implica observar las dimensiones relevantes, establecer el valor de la realidad evaluada en cada una de ellas y su importancia relativa, para finalmente combinarlos en forma ponderada en una conclusión o afirmación evaluativa. Scriven propone como ejemplo la evaluación de un mango en un puesto de frutas. Las principales propiedades a ser coligadas -los criterios de valor- son el color, la textura, el peso, el olor y la firmeza.

"El comprador ha aprendido cómo evaluar mangos por sí mismo o con el grupo de consumidores para el cual trabaja, esto es, (1) qué propiedades son relevantes (y deben ser coligadas); (2) cómo estimar cada una de ellas por la vista, el tacto y el aroma; (3) cómo la escala de valoración se vincula con la gama de colores, esencias o firmezas que pueden ser encontradas en la realidad, esto es, la escala de valoración". A partir de este ejemplo Scriven destaca un par de detalles importantes. En primer lugar, que algunas propiedades son evaluadas en forma dicotómica (existen o no existen) (por ejemplo, el olor); otras son evaluadas en una gradación (malo, regular, bueno, muy bueno) (por ejemplo, la firmeza); en tanto que otras pueden tener un valor numérico (por ejemplo, el peso). En segundo lugar, que algunas propiedades pueden tener mayor importancia que otras y constituir un punto de corte. Por ejemplo, si el mango es tan liviano que su interior parece ser una gelatina (firmeza), puede ser evaluado como inaceptable, aunque el color y el aroma sean muy buenos. En toda evaluación el evaluador "pondera" de algún modo las diferentes propiedades o aspectos relevantes de la realidad evaluada.

La evaluación en la educación se rige por esta misma lógica:

- tenemos un propósito: queremos decidir si estamos logrando lo que nos proponemos con nuestro curso, o cómo ayudar a los estudiantes a avanzar en su aprendizaje, o si un estudiante en concreto está o no en condiciones de aprobar el curso;
- necesitamos establecer una afirmación valorativa sobre la situación: cómo está el grupo en general o estudiantes específicos con relación a los aprendizajes que queremos promover, de modo que podamos definir los cursos de acción posibles, o, en otros casos, establecer las calificaciones de los estudiantes, que constituyen un juicio de valor sobre su desempeño;

- tomamos en consideración ciertas dimensiones o "criterios de valor" relevantes: nuestras definiciones de qué conocimientos, desempeños, habilidades y/o actitudes esperamos de los estudiantes como resultado del curso y qué importancia relativa tiene cada una de ellas;
- recogemos "evidencia empírica" sobre lo que cada estudiante puede hacer en cada una de las dimensiones o aspectos que definimos como relevantes: esto lo hacemos a través de tareas, pruebas, proyectos o de la observación directa del trabajo individual o en grupo, o del esfuerzo realizado por cada estudiante, entre otras muchas posibilidades;
- a partir del análisis de la "evidencia empírica" sobre los desempeños de cada estudiante en cada aspecto relevante, realizamos un trabajo de "coligación" bastante complejo, a través del cual establecemos una "conclusión evaluativa": en algunos casos será una calificación, en tanto que en otros será una valoración de la situación del grupo o de un estudiante en concreto y de los posibles cursos de acción para ayudar a los estudiantes a avanzar en su aprendizaje.

Al igual que en el ejemplo de los mangos, en general establecemos ponderaciones de los distintos contenidos y saberes que enseñamos. No todos tienen la misma importancia. Algunos pueden constituir un "punto de corte": *"si no sabe esto no puede aprobar el curso"*.

En el caso del mango, afirma Scriven, la evaluación es principalmente perceptiva: las valoraciones y su coligación surgen en forma inmediata de la observación. Algo parecido ocurre en muchas situaciones de la vida cotidiana, al igual que en ciertas disciplinas artísticas y deportivas. Por ejemplo, la gimnasia olímpica, los saltos ornamentales en piscina, la danza, la actuación de conjuntos de carnaval, la música, el teatro o la cata de vinos. En todos estos casos los evaluadores acuerdan explícitamente el conjunto de aspectos o dimensiones a valorar, y la valoración de las mismas emerge en forma inmediata de la percepción. La "evidencia empírica" es la observación. Por ejemplo, en el caso de un conjunto de carnaval, los miembros del jurado evalúan la letra, la música y el vestuario, entre otros aspectos. Si bien para el observador externo esta evaluación puede parecer arbitraria o caprichosa, en realidad está basada en una importante acumulación de experiencia de cada miembro del jurado. No cualquiera puede evaluar vinos, ni gimnasia olímpica ni danzas. Pero, al mismo tiempo, como es sabido que la evaluación no es "objetiva", sino que interviene la apreciación subjetiva del evaluador, se recurre a un jurado integrado por varios miembros, como forma de neutralizar el inevitable sesgo de cada uno de ellos. Es importante, además, retener lo seña-

lado unas líneas más arriba: si bien la valoración es en cierto grado subjetiva porque se desprende de la percepción directa del evaluador, las dimensiones a observar y los "criterios de valor" son explícitos, acordados previamente y conocidos por todos, evaluadores y evaluados.

En otros casos la evaluación no es perceptiva sino inferencial. Por ejemplo, en la evaluación que antecede a la determinación de comprar o alquilar una vivienda o a la decisión de cómo mejorar un programa para la erradicación del trabajo infantil. En estas situaciones la importancia de definir los referentes valorativos o criterios de valor es similar a los casos de evaluación perceptiva antes analizados, pero la cuestión de la evidencia empírica es más complicada. No alcanza con la observación directa, sino que se requiere de un trabajo más o menos extenso de recopilación y análisis de información, a partir de lo cual será necesario realizar inferencias. Por ejemplo, a partir de la información sobre cinco posibles departamentos para alquilar, del análisis de la misma y de la ponderación de las principales "dimensiones" (la ubicación, la cantidad de dormitorios, el estilo, la iluminación y el costo mensual), llegamos a una conclusión evaluativa sobre cuál es la mejor opción para nosotros en este momento. Al igual que en el caso de la evaluación perceptiva, la inferencial no da lugar a una única conclusión evaluativa posible. Diversas personas pueden llegar a valoraciones diferentes a partir del mismo conjunto de información.

En la educación en general realizamos ambos tipos de evaluaciones. Muchas veces operamos en forma perceptiva: a partir de la simple observación directa de la actuación de un estudiante o de la revisión de su trabajo, llegamos directamente a una valoración. Otras veces actuamos en forma inferencial: analizamos varias piezas de información sobre el desempeño del estudiante para llegar a una conclusión.

Al igual que en los ejemplos antes planteados (la cata de los vinos o los conjuntos de carnaval), la evaluación en educación descansa fuertemente en la acumulación de saber y de experiencia de cada docente. Y, del mismo modo, el resultado de una evaluación, tomando como punto de partida los mismos trabajos de los mismos estudiantes, puede diferir en función del docente que evalúa. Esto no necesariamente es malo en sí mismo. Cierto grado de subjetividad es parte de la lógica de la evaluación en todos los campos de la actividad humana. El problema no es la subjetividad, sino la opacidad: no siempre explicitamos y comunicamos con claridad cuáles son los aspectos o dimensiones importantes, ni sobre la base de qué criterios y con qué evidencias estamos evaluando.

Los criterios que utiliza un jurado de gimnasia olímpica son explícitos. Los atletas los conocen. Los criterios que utiliza un catador de vinos son

compartidos entre los entendidos, pero opacos para los no especializados. Cuando uno lee en la etiqueta de un vino que tiene *"un intenso gusto a fruta madura con notas de cacao y café y un sutil aroma a madera"*, lo prueba y no sabe de qué le están hablando, salvo que sea un consumidor con cierta "especialización". Algo parecido le pasa a los estudiantes con muchas de nuestras evaluaciones: no comprenden mucho de lo que les decimos cuando los evaluamos. Para ellos somos como un sommelier (de allí que se fijen solamente en la nota). Una de las principales dificultades en la evaluación en educación es que, en general, no tenemos un abordaje sistemático, claro y explícito de la misma y de su lógica. Tendemos a pensar la evaluación a partir de ciertos procedimientos, preguntas, actividades o tipos de ejercicios que hemos ido elaborando a lo largo de los años, a partir de la experiencia. Nos enfocamos mucho en los instrumentos (la evidencia) y poco en los criterios o referentes, que no solemos explicar suficientemente a nuestros estudiantes. Según analizaremos en profundidad a lo largo del libro, la cuestión de la comunicación de qué aprendizajes queremos que logren nuestros estudiantes y qué es un desempeño aceptable o uno destacado, es uno de los desafíos más importantes para la evaluación en el aula.

Una segunda cuestión importante que analizaremos en este libro es la relativa a las finalidades de la evaluación y a la distinción entre evaluación para la certificación y evaluación para el aprendizaje. El modo de llevar adelante la evaluación depende fuertemente de su finalidad o propósito. En educación evaluamos con dos finalidades principales: para dar cuenta en forma pública de lo que han logrado los estudiantes (para lo cual establecemos calificaciones) y para ayudarles a avanzar en sus procesos de aprendizajes (para lo cual necesitamos propiciar instancias de evaluación formativa). Según analizaremos más adelante, estas dos finalidades dan lugar a dos formas de evaluación completamente diferentes. Sin embargo, los docentes en general tendemos a no diferenciarlas y a calificar a los estudiantes con excesiva frecuencia. En los capítulos 4 y 5 explicaremos por qué ocurre esto, las diferencias fundamentales entre ambas formas de evaluación y cómo realizar mejor cada una de ellas.

Un tercer desafío fundamental es aprender a actuar más como entrenadores que como jurados. Un entrenador de gimnasia olímpica que continuamente le diese un puntaje a sus atletas, similar al que le darán los jurados en la competencia, no los ayudaría mucho a mejorar su desempeño. El papel del entrenador no es dar puntajes, sino devoluciones que ayuden a que el atleta aprenda a percibir y valorar sus movimientos y "darse cuenta" de qué es lo que tiene que corregir. Nuestra situación como docentes es similar a la del entrenador de gimnasia olímpica: necesitamos desarrollar estrategias orien-

tadas a que nuestros estudiantes desarrollen su capacidad para apreciar qué es un buen desempeño o un buen trabajo en nuestra disciplina o materia. Necesitamos trabajar en el desarrollo de la capacidad de autopercepción y autoevaluación de los estudiantes. Trabajaremos este aspecto en profundidad en el capítulo 4.

3. La enseñanza como práctica cultural

La mayoría de los docentes intentamos, a lo largo de los años, ir cambiando y mejorando nuestras formas de enseñar y de evaluar. Algunas cosas logramos, pero muchas prácticas y rutinas permanecen incambiadas. Cuando participamos en un taller o leemos algo nuevo, muchas veces nos entusiasmamos con la idea de intentar nuevas propuestas de enseñanza y/o de evaluación. Pero la mayoría de las veces el entusiasmo queda rápidamente por el camino. El ambiente de trabajo, las dificultades, el aislamiento en el que trabajamos, las demandas de los propios estudiantes, nos retrotraen a las prácticas habituales y nos ganan las rutinas que tenemos construidas. Cambiar los modos de trabajar en el aula es complejo. ¿Por qué es tan difícil?

Una primera razón importante detrás de las dificultades para modificar las formas de enseñar está vinculada con las concepciones predominantes acerca del trabajo docente y las condiciones de trabajo que se derivan de dichas concepciones: la falta de reconocimiento al trabajo fuera del aula y el desempeño "en solitario".

Enseñar es una tarea difícil y exigente. Todos los días es necesario preparar y llevar adelante una o varias "representaciones" o performances, frente a uno o varios grupos de niños, niñas y/o adolescentes. Estas performances incluyen explicaciones y actividades que interesen y, ojalá, entusiasmen a los estudiantes, en torno a un conjunto de saberes que la sociedad y los educadores hemos decidido que son importantes. Intentamos interesarlos con las ciencias, la literatura, la historia y las artes. Nos proponemos, por encima de todo, contribuir al desarrollo personal y social de nuestros estudiantes, a que aprendan a vivir con otros, a concebir y llevar adelante proyectos, a participar en la comunidad y en la sociedad de las que forman parte. Cada día, cada docente, debe realizar una o varias performances distintas. En primaria es una y dura varias horas. En secundaria es más breve, pero el docente debe realizar varias diferentes para distintos grupos de estudiantes.

El día a día en el aula requiere en forma continua la preparación de presentaciones y exposiciones; la concepción de nuevas actividades; la revisión y modificación de lo realizado en oportunidades anteriores; la búsqueda de nuevos materiales; la elaboración de actividades específicas para situaciones

específicas o estudiantes concretos. Y, además de lo anterior, revisar y corregir las tareas realizadas por los estudiantes.

El problema central es que para hacer esto bien se necesita tiempo de preparación, en lo posible con otros colegas. Quien haya dado alguna vez una clase o una conferencia sabe que cada hora de "actividad en el aula" requiere, como mínimo, otra hora para prepararla. Sin embargo, en la mayor parte de los países de la región el trabajo docente es concebido y remunerado como horas en el aula[2]. Y, en parte como consecuencia de ello, la docencia es concebida como una actividad estrictamente individual.

Con las distancias del caso, la situación es similar a la de un actor de teatro que tuviese que representar un nuevo guión cada noche, sin tiempo para ensayar y, además, sin director ni colegas de los cuales recibir sugerencias y orientaciones. Por otra parte, el actor solo debe preocuparse por su actuación. Debe representar bien su papel, pero no es responsable de que los espectadores comprendan la obra. Es más, cada espectador puede entender algo distinto y eso no constituye un problema. El docente, en cambio, debe asegurarse de que cada estudiante comprenda aquello que quiso trasmitir a sus alumnos.

En primaria el maestro tiene la ventaja de trabajar con menos grupos de alumnos (normalmente uno o dos), pero tiene la desventaja de que debe preparar clases sobre todas las áreas y temas (lenguaje, matemáticas, historia, geografía, formación ciudadana...). El profesor de educación media tiene la ventaja de concentrarse en un área de conocimiento o disciplina, pero la desventaja de trabajar con varios grupos distintos de estudiantes y, muchas veces, con diversidad de cursos para distintos niveles.

Además de lo anterior, la docencia es una actividad predominantemente solitaria. Si bien uno interactúa con colegas todos los días, lo que ocurre

---

2. De acuerdo con el estudio TALIS sobre la profesión docente en 34 países (realizado en 2013 y publicado por la OCDE en 2014) la dedicación promedio de los profesores en la enseñanza media es de 38 horas semanales. La mitad de ellas (19 horas semanales en promedio) son dedicadas a trabajar en el aula enseñando, mientras que la otra mitad se destina a la realización de actividades fuera del aula. Los países en los que los profesores destinan menos horas a la docencia directa son Noruega (15 horas) y Japón (18 horas). La consecuencia obvia es que en estos países los docentes disponen de más tiempo para preparar sus clases, evaluar las tareas de sus estudiantes, trabajar en equipo y otras actividades. De acuerdo con este estudio, los profesores dedican, en promedio, siete horas semanales a preparar sus clases, cinco a corregir y evaluar trabajos de sus estudiantes y dos a actividades extracurriculares y encuentros con las familias. En la mayor parte de los países de la OCDE estas horas son parte del trabajo remunerado del profesor.

dentro del aula es, en cierto modo, "secreto". No porque nadie sepa cómo trabaja cada uno, sino porque de eso en general no se habla. En realidad es una suerte de "secreto a voces": aunque todos en una institución educativa sepan qué dificultades tiene cada docente dentro del aula, no se comparten ni son objeto de intercambio y análisis. En nuestros países, el trabajo en equipo que involucre la observación y el análisis de las formas de llevar adelante las clases son muy poco comunes. Por otra parte, las visitas de supervisores suelen tener un carácter predominantemente formal y administrativo.

Una de las consecuencias principales de lo anterior es la dificultad para crear y compartir un "corpus" de conocimiento profesional. Muchos docentes desarrollan importantes e interesantes experiencias de enseñanza. A lo largo de este libro tendremos la oportunidad de presentar varios casos de experiencias de evaluación desarrolladas por docentes. Pero estas experiencias son poco conocidas por sus colegas. A lo sumo se comparten entre los dos o tres más cercanos. El conocimiento eminentemente práctico y aplicado, generado en el marco de la actividad docente, tiene muy poca circulación y casi no se acumula. Los docentes tendemos a leer lo que escriben sobre la educación y la enseñanza otros especialistas que, como los autores de este libro[3], observan el trabajo docente desde fuera del aula.

La falta de tiempo para la preparación de clases y el aislamiento en que se desarrolla la labor docente son, por tanto, dos limitaciones centrales para la innovación en la educación y para la creación y acumulación de conocimiento desde y sobre el trabajo de la propia profesión. La consecuencia principal es que cada docente va construyendo, a lo largo de los años, un conjunto de prácticas y modos de hacer las cosas, que le permite llevar el día a día y "sobrevivir" a las demandas cotidianas del aula, que se van consolidando y que son muy difíciles de modificar.

Pero existe una razón más profunda detrás de las dificultades para cambiar las formas de hacer las cosas en el aula: la enseñanza es, ante todo, una actividad cultural. Enseñar *"es más parecido a participar en una cena familiar que a aprender a usar una computadora...* [en el sentido de que no se aprende como una técnica en un momento dado, sino por acumulación de experiencia de manera informal a través de largos períodos de tiempo] *...mucha gente cree que la enseñanza es una habilidad innata, algo con lo que uno nace. Otros creen que los docentes aprenden a enseñar haciendo una carrera en los Institutos de Formación Docente. Nosotros creemos que ninguna de estas descripciones es la*

3. Si bien nos consideramos parte del cuerpo docente, somos plenamente conscientes de que nuestra actividad y nuestra mirada sobre la educación, en este momento, es básicamente externa a la vida cotidiana en las aulas.

más apropiada. A enseñar, al igual que otras actividades culturales, se aprende a través de la participación directa durante largos períodos de tiempo. Es algo que uno aprende más por crecer en una determinada cultura que por estudiarlo formalmente... La mayoría de las personas no estudió para ser docente pero han sido estudiantes. Por eso la gente dentro de una determinada cultura comparte una imagen mental de qué es enseñar. Llamamos a esta imagen mental "guión" ["*script*", en inglés, en el original]... *creemos que los guiones explican por qué las clases dentro de un país siguen un patrón común, a pesar de las diferencias individuales: las clases fueron preparadas y enseñadas por docentes que comparten los mismos guiones"* (Stigler y Hiebert, 1999: 86-87).

Estos autores realizaron varios trabajos comparando las formas de enseñar matemáticas y ciencias en distintos países del mundo, basándose en cientos de clases filmadas en video (Hiebert y otros, 2003; Roth y otros, 2006). Quedaron sorprendidos por lo mucho que varían las formas de enseñar entre países y lo mucho que se parecen dentro de un mismo país (Stigler y Hiebert, 1999: 11).

De acuerdo con los autores, en los Estados Unidos los profesores conciben la matemática como un conjunto de procedimientos que sirven para resolver problemas. También creen que aprender términos nuevos y practicar procedimientos no es interesante, por lo cual buscan hacer la clase entretenida haciendo bromas o conversando un rato de otros temas. Su objetivo principal es que los alumnos aprendan destrezas para resolver ejercicios. En Japón, en cambio, los profesores conciben la matemática como un conjunto de relaciones entre conceptos, hechos y procedimientos. Estas relaciones se descubren mediante el desarrollo de métodos variados de resolución de problemas nuevos, estudiando dichos métodos, buscando los más efectivos y discutiendo las propiedades que subyacen a cada uno. Los profesores japoneses actúan como si la matemática fuese naturalmente interesante y partiendo de la base de que a los alumnos les resultará interesante explorar soluciones a los problemas. Su objetivo principal es que los alumnos aprendan a pensar sobre las cosas de maneras nuevas. En Japón, dentro del aula, los profesores no suelen hacer bromas ni conversar sobre temas ajenos a la clase.

Los autores ilustran estas diferencias a través de dos formas típicas de enseñar operaciones con fracciones. En los Estados Unidos los profesores comienzan trabajando con las fracciones más simples (por ejemplo, 1/5 + 2/5), luego pasan a actividades un poco más complejas (por ejemplo, 1/2 + 1/4), para recién al final llegar a lo más difícil. La idea es ir complejizando progresivamente la propuesta, graduando las dificultades y practicando mucho. En Japón, en cambio, los profesores comienzan planteando directamente una operación difícil (por ejemplo, 2/3 + 4/7). Creen que los estudiantes apren-

den mejor intentando resolver problemas por sí mismos y discutiendo las posibles soluciones. La frustración y la confusión se consideran parte natural del proceso. Esto requiere tiempo para explorar e inventar, para equivocarse, para recibir orientación en el momento oportuno y para reflexionar. El propio manual del maestro dice que los estudiantes comprenderán mejor si se equivocan y examinan las consecuencias de ello.

De acuerdo con Hiebert y Stigler (1999), las actividades mediante las cuales los docentes interactúan con sus alumnos no están determinadas principalmente por su formación, ni se modifican a través de la misma. Las prácticas no se inventan, se heredan de las generaciones anteriores. "*La mayor parte de lo que ocurre en el aula está determinado por un código cultural que funciona, en algunos sentidos, como el ADN de la enseñanza*" (prefacio, p. xii). Esto ocurre, en buena medida, porque aprendemos a enseñar desde niños, cuando vamos a la escuela y, antes aún, cuando jugamos a estar en la escuela. Por esta razón los "guiones" comienzan a conformarse muy tempranamente, incluso antes de entrar a la escuela. Luego se consolidan a lo largo de 12 años de escolaridad.

El carácter cultural de la enseñanza explica las dificultades para cambiarla, dado que tanto los docentes como los estudiantes comparten los mismos "guiones" y saben qué deben esperar unos de otros y qué roles deben desempeñar (Hiebert y Stigler, 1999). Las reflexiones anteriores incluyen, por cierto, a la evaluación. Los modos en que evaluamos a nuestros estudiantes, en particular las prácticas vinculadas con la asignación de calificaciones, forman parte del ADN del modelo educativo en el que trabajamos.

Si queremos introducir cambios en las prácticas de enseñanza y evaluación es necesario moverse de la visión de la enseñanza como actividad solitaria, propia de cada uno, hacia una visión de la enseñanza como una actividad profesional abierta a la observación, el estudio y el mejoramiento colectivos. Las tradiciones instaladas en la cultura del sistema educativo son muy difíciles de modificar si no existe una práctica sistemática de revisión e intercambio entre docentes en torno a ellas. *"Es necesario darse cuenta de cuáles son las rutinas culturales que gobiernan la vida en el aula, cuestionar los supuestos de estas rutinas, y trabajar para modificarlas durante un largo tiempo. Implica reconocer que los detalles de lo que los docentes hacen -las preguntas específicas que formulan, el tipo de tareas que proponen a los estudiantes, las explicaciones que dan- son cosas fundamentales para el aprendizaje de los estudiantes..."* (101). Para ello es preciso abrir las puertas del aula, invitar a otros a observar y criticar, no para imitar, sino para participar de la gratificante tarea de mejorar las prácticas de toda la profesión docente.

Crear las condiciones para la construcción de una visión profesional y colectiva sobre la enseñanza, centrada en el análisis y la mejora de las prácticas es responsabilidad fundamental de autoridades, políticos y sindicatos. Mientras tanto, seguiremos haciendo reformas curriculares sin que cambie mayormente lo que acontece dentro de las aulas.

En este libro les proponemos comenzar a hacer de la enseñanza y de la evaluación objetos de discusión, análisis y experimentación. Las propuestas de evaluación que empleamos en el aula son, en primer lugar, un buen indicador de qué y cómo enseñamos. Y, al mismo tiempo, un buen punto de partida para iniciar un proceso colectivo de revisión y cambio en las prácticas. Compartir y analizar junto con otros colegas nuestras formas de evaluar a los estudiantes es un excelente y concreto camino para iniciar un diálogo enriquecedor. Requiere una cierta dosis de humildad. La invitación que hacemos a los lectores de este libro es poner en cuestión los modos en que llevan adelante las actividades en el aula, a partir de una mirada reflexiva sobre sus propias evaluaciones.

4. La evaluación de aprendizajes en las aulas de América Latina

Tal como explicamos en la presentación del libro, las ideas y propuestas incluidas en el mismo tienen como una de sus fuentes principales dos trabajos de investigación que realizamos en los años 2008 y 2012 en nueve países de la región[4]. Ambos estudios estuvieron motivados por el propósito de conocer mejor y “abrir ventanas” hacia el interior de las aulas, que permitiesen observar y analizar las prácticas de evaluación de aprendizajes que tienen lugar en ellas, bajo el supuesto de que estas tienen una importancia crucial para la experiencia educativa y el aprendizaje de los estudiantes.

El primero de los estudios, denominado **“La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas”**[5], incluyó grupos y maestros de 6º grado de primaria. Se analizaron las prácticas de evaluación en Lenguaje y Matemática. Incluyó los siguientes países:

4. En rigor, en nueve ciudades: Ciudades de Guatemala y México, San Salvador, Buenos Aires, Cali, Lima, Montevideo, San José de Costa Rica y Santiago de Chile.

5. Contó con el apoyo financiero del Grupo de Trabajo sobre Estándares y Evaluación de PREAL (GTEE-PREAL).

Uruguay, Perú, El Salvador, Guatemala, Costa Rica, Colombia, Argentina y México[6].

El segundo estudio, denominado **"La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina"**[7] se enfocó en 3º de secundaria básica y en las materias vinculadas a las Ciencias Naturales: Biología, Química y Física. Se realizó en Uruguay, Perú, Chile y Colombia.

Los estudios estuvieron orientados por las siguientes interrogantes:

- ¿Cuáles son los enfoques y concepciones en las que se basan las prácticas de evaluación de los docentes?
- ¿Qué tipo de tareas, preguntas o consignas utilizan para evaluar aprendizajes?
- ¿Cómo emplean las evaluaciones para orientar a sus estudiantes y ayudarles a progresar?
- ¿Cómo comunican los resultados de las evaluaciones?
- ¿Qué criterios priorizan al calificar a sus estudiantes y al tomar decisiones de promoción o reprobación?
- ¿Qué semejanzas y diferencias hay entre los países estudiados?

Ambos fueron estudios de casos, por lo cual sus resultados no tuvieron la pretensión de ser generalizables a cada país. Se trató de estudios de carácter exploratorio, dirigidos a recoger información sobre las concepciones y propuestas de evaluación de los docentes. Para cada estudio se seleccionaron diez centros educativos del nivel correspondiente, en cada país. La selección, que fue de carácter intencional, estuvo orientada a establecimientos de la ciudad en la que se realizó el estudio, que estuviesen ubicados preferentemente en sectores populares, pero incluyendo algunos de contextos socioculturales más favorecidos. Se estableció como criterio que los centros educativos seleccionados debían tener resultados educativos en el promedio del país o por encima. El propósito de este criterio fue trabajar con un conjunto de docentes que, en cada país, reflejase lo que hace un buen docente y las formas en que son evaluados los niños de sectores populares por dichos docentes. Deliberadamente se evitó ir a establecimientos en los que las condiciones fuesen extremadamente difíciles, porque en esos casos suelen verificarse problemas

6. Inicialmente el estudio fue concebido para ser realizado en Uruguay, El Salvador, Guatemala, Costa Rica y Perú. Luego, a partir de la iniciativa de instituciones locales mencionadas en la Presentación de este libro, se incorporaron Argentina, Colombia y México.

7. Se realizó en el marco del Fondo concursable para proyectos de investigación (FIE-PREAL).

graves que impiden que la enseñanza se desarrolle con normalidad. Se evitó también incluir centros educativos de los sectores más favorecidos de la sociedad, para evitar situaciones atípicas.

Se incluyó una mayoría de centros públicos y algunos privados. Los establecimientos fueron contactados a través de una institución colaboradora en cada país y su participación fue de carácter voluntario, al igual que la de los profesores. En cada centro educativo se trabajó con dos docentes, es decir, un total de veinte por cada país.

Las dos investigaciones fueron de carácter cualitativo y tuvieron dos ejes principales para la recogida y análisis de información: a) la realización de una entrevista en profundidad de alrededor de 45 minutos con cada docente, durante las cuales se conversó con los entrevistados acerca de sus ideas sobre la evaluación, sus criterios para evaluar, los instrumentos y procedimientos empleados y las formas de utilizar sus resultados; b) un relevamiento fotográfico de una muestra de las pruebas y otro tipo de actividades de evaluación empleadas por cada docente en sus clases, tanto con fines formativos como para calificar a sus estudiantes. Estas muestras fueron seleccionadas por el propio docente. Esto implica que, en caso de existir un sesgo en la selección de propuestas, el mismo es hacia lo que cada profesor considera lo mejor de su trabajo. En el primer estudio se aplicó un cuestionario autoadministrado a los docentes (enfocado en su experiencia con las evaluaciones externas) y en el segundo a los estudiantes (enfocado en sus experiencias y percepciones acerca del modo en que son evaluados en el aula). En ambos casos se hizo, además, una revisión de programas de estudio y normas de evaluación para los grados y disciplinas escogidos.

Como resultado del relevamiento efectuado en el primer estudio, se conformó un banco de archivos de audio digital que incluye 158 entrevistas a maestros y 4360 registros fotográficos de propuestas de evaluación y trabajos de estudiantes. En el segundo estudio el banco de registros incluyó 76 entrevistas en profundidad a profesores de ciencias; una colección de 353 pruebas y propuestas de evaluación que contienen un total de 4359 tareas; y 1824 cuestionarios respondidos por estudiantes.

A lo largo de los capítulos que siguen emplearemos muchos de estos materiales para ilustrar las formas en que los docentes de la región conciben la evaluación y la llevan adelante en sus aulas. Retomaremos también algunos de los análisis realizados en varios de los productos generados a partir de estos estudios (Loureiro 2009; Picaroni 2009; Ravela 2009; Ravela y otros 2104).

Referencias bibliográficas

Hiebert, J. y otros (2003). *Teaching Mathematics in Seven Countries. Results From the TIMSS 1999 Video Study*. Washington D.C.: National Center for Education Statistics (NCES).

Loureiro, G. (2009). *Evaluación en el aula, currículo y evaluaciones externas*. Lima: GTEE-PREAL.

Picaroni, B. (2009). *La evaluación en las aulas de primaria: usos formativos, calificaciones y comunicación con los padres*. Lima: GTEE-PREAL.

Ravela, P. (2003). *Los elementos centrales de los procesos de evaluación educativa*. Conferencia presentada en el VII Congreso Nacional de Educadores, realizado en Lima en agosto de 2003. Disponible en: https://www.academia.edu/8659759/Los_elementos_centrales_de_los_procesos_de_evaluación_educativa Acceso: 17/01/17.

Ravela, P. (2006). *Fichas Didácticas. Para comprender las evaluaciones educativas*. Santiago de Chile: PREAL - Editorial San Marino.

Ravela, P. (2009). *Consignas, devoluciones y calificaciones: los problemas de la evaluación en las aulas de educación primaria en América Latina*. En: Páginas de Educación, Vol. 2, pp. 49-89. Montevideo: Universidad Católica del Uruguay.

Ravela, P. y otros (2014). *La evaluación en las aulas de secundaria básica en cuatro países de América Latina*. En: Propuesta Educativa, Año 23 Nº 41, pp. 20 a 45. Buenos Aires: FLACSO.

Rivas, A. (2015). *América Latina después de PISA: lecciones aprendidas de la educación en siete países 2000-2015*. Buenos Aires: Fundación CIPPEC.

Roth, K.J. y otros (2006). *Teaching Science in Five Countries: Results From the TIMSS 1999 Video Study*. Washington, DC: National Center for Education Statistics (NCES).

Scriven, M. (2013); *The Foundation and Future of Evaluation*. En: Donaldson, S. (editor); The Future of Evaluation in Society. A tribute to Michael Scriven. Information Age Publishing.

Stigler, J. & Hiebert, J. (2009). *The Teaching Gap*. New York: Free Press.

CAPÍTULO 2

¿QUÉ APRENDIZAJES PROMUEVEN LAS ACTIVIDADES DE EVALUACIÓN QUE PROPONEMOS A NUESTROS ESTUDIANTES?

Introducción

La mayoría de los docentes querríamos que nuestros estudiantes se motiven con el aprendizaje y con todo aquello que intentamos enseñarles. Querríamos que no se limiten a memorizar y repetir los temas que trabajamos con ellos en el aula, sino que los "entiendan" y sean capaces de reflexionar críticamente en torno a dichos temas, que se apropien de los contenidos y que sean capaces de seguir aprendiendo a partir de ellos. Querríamos facilitarles el acceso a saberes que les permitan comprender mejor la sociedad en la que viven y les habiliten para llevar adelante sus proyectos personales. Querríamos contribuir a la formación de ciudadanos reflexivos y creativos, capaces de participar activamente en sus comunidades y en la sociedad. Este tipo de propósitos educativos suelen estar expresados de una u otra forma en las leyes de educación de los países y en los currículos nacionales. Prácticamente ningún docente discrepa con ellos, antes bien, casi todos querríamos que nuestras clases sirvieran a los mismos.

Y sin embargo, estos propósitos son escurridizos, difíciles de lograr en la práctica. Muy a nuestro pesar, muchas veces apenas logramos plasmarlos en nuestras clases. La mayoría tendemos a destinar la mayor parte del tiempo a hablar a nuestros estudiantes sobre los temas del programa y a intentar involucrarlos en actividades que se supone les ayudarán a aprender. A veces exponemos o explicamos en el pizarrón, a veces los hacemos trabajar en equipos, a veces les pedimos que resuelvan ejercicios, que lean un texto o que discutan un tema. A veces logramos que la mayoría hagan las tareas (con más o menos "ganas") y que algunos, los menos, se entusiasmen. Otras veces notamos que la mayoría no sintoniza con la propuesta, que no se involucra

realmente con actividades que proponemos, que las realizan para cumplir con nosotros o simplemente por obligación. En algunas ocasiones nos vemos en la necesidad de emplear recursos tales como "la nota" o alguna advertencia más o menos explícita vinculada con la aprobación del curso, con el fin de "motivarlos".

La mayor parte del tiempo, el relacionamiento que logramos entre los estudiantes y el contenido que enseñamos suele ser más bien superficial. Buena parte de ellos no logra comprender del todo lo que está estudiando y se queda con algunas ideas elementales, poco conectadas entre sí, que puede recordar y repetir en forma aislada para responder preguntas muy concretas. Muy pocos logran construir una conceptualización del tema que les permita explicarlo con sus propias palabras, desarrollar una reflexión propia o resolver situaciones nuevas en las que dichos temas o conceptos están involucrados. Como consecuencia, solo llegamos a motivar a algunos pocos estudiantes, los más predispuestos al trabajo intelectual y abstracto o los que tienen alguna inclinación natural hacia nuestra materia. Pero la mayoría se "conecta" con nuestras clases de manera intermitente, intenta "cumplir" con ciertos mínimos que nos dejen satisfechos a nosotros y que les permitan a ellos promover el curso.

Por diversas causas de índole social y cultural, vinculadas principalmente con los nuevos medios y dispositivos de comunicación existentes, con los que los niños se familiarizan desde muy pequeños, la disciplina necesaria para el trabajo escolar está mucho menos extendida que algunas décadas atrás. Estar sentado durante un tiempo relativamente prolongado, seguir con atención una explicación, escuchar y leer, concentrarse en la realización de una tarea e, inclusive, trabajar durante un cierto período de tiempo interactuando con compañeros, no son hábitos que puedan esperarse como pre-existentes. Este fenómeno se produce, con algunos matices, tanto en primaria como en secundaria. Los niños más chicos suelen estar más predispuestos a aceptar lo que les propone su maestro o la escuela, en tanto que los adolescentes suelen ser más reticentes y difíciles de motivar.

Como resultado de estas situaciones los docentes en general nos sentimos insatisfechos con nosotros mismos y con nuestro trabajo. Tendemos a construir justificaciones que colocan el problema "afuera" de nosotros: las familias no apoyan, no los motivan, las nuevas generaciones no tienen hábitos de estudio, se pasan conectados a redes sociales, el programa y la inspección nos obligan a dar todos estos temas...

Cada lector se sentirá más o menos reflejado en las descripciones anteriores. Es probable que la experiencia de algunos sea diferente. En cualquier caso, la invitación que hacemos en este capítulo es a revisar las formas en que

presentamos el "contenido" a nuestros estudiantes, tomando como punto de partida la revisión de las actividades de evaluación que les proponemos. Estas son un reflejo de lo que enseñamos y comunican a los estudiantes y a las familias qué es lo que en la práctica valoramos como aprendizaje. Hablan por sí mismas, mucho más que cualquier discurso sobre intenciones educativas. Observando las tareas de evaluación podemos concluir mucho sobre qué enfoque estamos dando a los contenidos curriculares y en qué tipo de experiencias de aprendizaje estamos involucrando a nuestros estudiantes: "*dime cómo evalúas y te diré cómo enseñas*" (Picaroni y Loureiro 2010).

> "*La evaluación cumple un rol muy importante dentro del proceso de enseñar y de aprender y el uso apropiado de la información que de ella se deriva es fundamental para la mejora de los aprendizajes. El propio desarrollo de la enseñanza necesita de la evaluación formativa. Ella actúa como proceso regulador entre la acción del docente y el aprendizaje del alumno, ya que permite ajustar las intervenciones del maestro a las necesidades del que aprende. También en la evaluación sumativa y para calificar se pone de manifiesto qué es lo que los docentes priorizan en el aprendizaje de los alumnos. Por lo expuesto, las prácticas de evaluación ofrecen un insumo muy adecuado para conocer las prácticas de enseñanza...*" (Picaroni y Loureiro 2010: 31).

Si las actividades de evaluación que realizan nuestros estudiantes están enfocadas principalmente en la memorización y aplicación rutinaria de ciertos conocimientos abstractos, difícilmente estaremos formando estudiantes y ciudadanos reflexivos y críticos. Y, sobre todo, tal vez se encuentre allí una de las raíces de la falta de interés y motivación de nuestros estudiantes (y de buena parte de nuestra insatisfacción con nuestro trabajo).

El propósito de este capítulo es ofrecer a los lectores un conjunto de claves para analizar, en forma individual o colectiva -ojalá esta última-, las actividades de evaluación que proponen habitualmente a sus estudiantes, el grado en que son consistentes con sus intenciones educativas y la forma de mejorarlas.

1. Conceptos clave para el análisis de actividades de evaluación

A lo largo del libro emplearemos los siguientes términos para caracterizar las actividades de evaluación. Denominaremos "*propuesta*" a una actividad -o varias- que conforman una instancia de evaluación como, por ejemplo, una prueba, un examen, un escrito, un parcial o a una instancia

de evaluación con fines formativos como un taller o la exhibición de un desempeño (en particular, aunque no solamente, en arte o educación física). *Propuesta de evaluación,* por tanto, es una expresión que utilizamos en forma genérica para referirnos a instancias de evaluación que pueden incluir una o varias actividades.

Una "*actividad*" de evaluación constituye una unidad en sí misma y normalmente contiene dos tipos de elementos principales:

1. *Consigna o enunciado*: toda actividad de evaluación incluye una expresión que le solicita al estudiante una tarea, "hacer algo": dar una respuesta, resolver un ejercicio o problema[1] para "hallar" la "solución", llevar adelante un desempeño, elaborar un texto o alguna otra forma de producción intelectual, escribir un ensayo sobre un tema, construir una propuesta, seleccionar la respuesta correcta de entre varias posibles, o construir una explicación, argumentación o reflexión que debe ser expresada oralmente o por escrito. Dicho en otras palabras, la consigna es la expresión de lo que se espera que el estudiante elabore en el marco de la actividad de evaluación planteada. En general se expresan por escrito, pero también puede haber consignas que se expresen verbalmente. Consigna y tarea son como dos caras de una moneda: la consigna es la formulación que propone el docente, la tarea es lo que el estudiante "hace".
2. Alguna forma de *información* que el estudiante debe "leer" y "comprender"[2]: un texto narrativo o informativo, un conjunto de datos, una tabla, gráfico o ilustración, entre otras posibilidades. Por lo general este conjunto de información constituye o está referida a lo que se denomina una *situación*[3], coloca al estudiante en un contexto en el que debe comprender y utilizar la información.

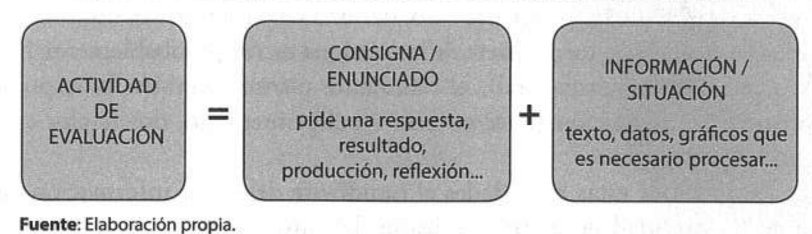
Según se mostrará enseguida, muchas actividades de evaluación contienen únicamente un enunciado o consigna, sin información ni situación a procesar (por ejemplo, ¿Quién fue el autor de "La Ilíada"?).

Una primera forma de aproximarse a la revisión de las prácticas de evaluación en el aula consiste en analizar las actividades que proponemos a

1. Que, como veremos más adelante, no son lo mismo, aunque se suele utilizar ambas expresiones como sinónimos.
2. Leer en sentido amplio, incluyendo la lectura de gráficos o la captación de la información incluida en una ilustración.
3. De acuerdo con el DRAE, una situación es un "conjunto de las realidades o circunstancias que se producen en un momento determinado y que determinan la existencia de las personas o de las cosas".

nuestros estudiantes desde esta doble perspectiva: i) ¿qué necesita "pensar" y "hacer" el estudiante para producir/construir lo que le pide la consigna que le hemos planteado?; y, ii) ¿qué tipo de información o situación debe leer, comprender y procesar?

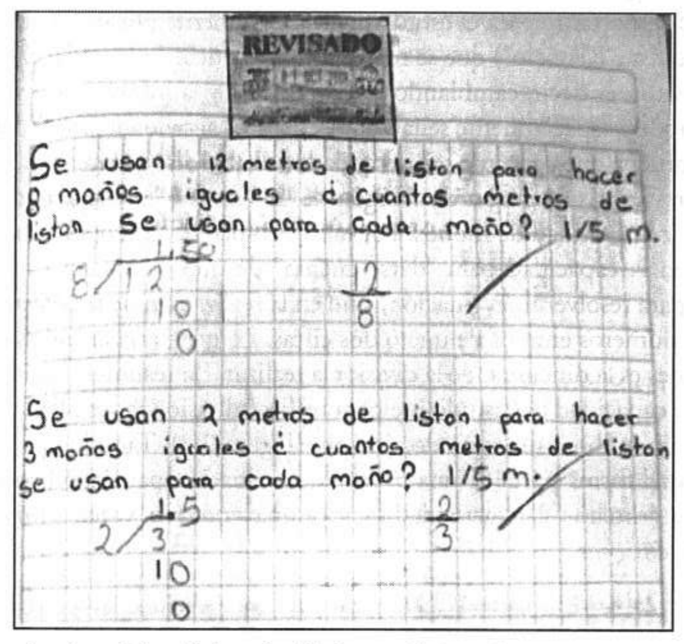
Figura 2.1: Elementos constitutivos de las actividades de evaluación



Fuente: Elaboración propia.

En la Figura 2.2 se pueden observar dos actividades de evaluación que forman parte de una prueba. La prueba tiene en total seis actividades, todas similares a las incluidas en la Figura.

Figura 2.2: Actividad de evaluación en Matemática en 6º de Primaria



Fuente: Base de registros del estudio **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas**.

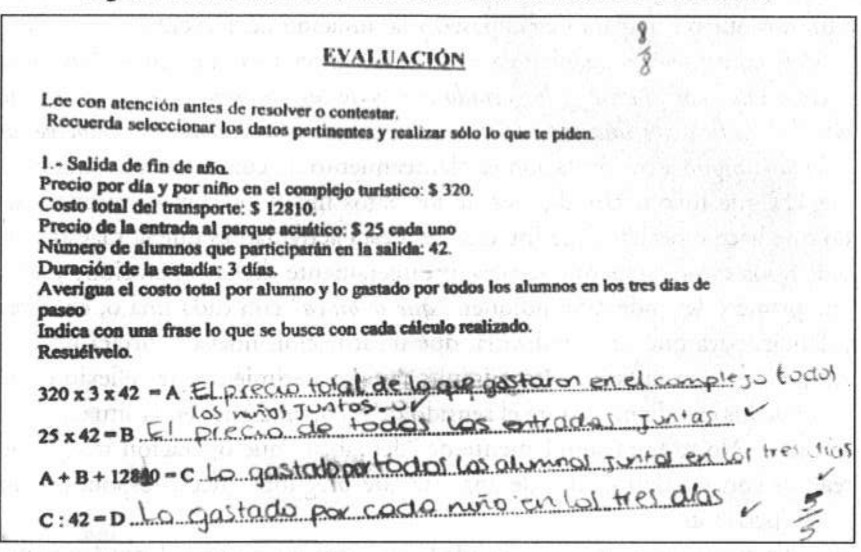
Como se puede observar, ambas actividades incluyen cierta ***información***: *"12 metros de listón y 8 moños iguales"* en la primera actividad; *"2 metros de listón y 3 moños iguales"* en la segunda. El texto de la actividad coloca a los estudiantes en una **situación**: *"hacer moños"*. Ambas actividades tienen el mismo **enunciado** o **consigna**, constituida por la pregunta *"¿Cuántos metros de listón se usan para cada moño?"*. Esta pregunta indica al estudiante qué es lo que se espera que haga. En este caso, dar una respuesta que es un resultado numérico. Si bien no forma parte de la consigna escrita (probablemente haya sido expresado en forma oral), el estudiante plantea también la respuesta expresada como fracción (doce octavos en el primer caso, dos tercios en el segundo).

Para realizar estas actividades el estudiante debe leer información elemental (la cantidad de metros de listón, la cantidad de moños y el hecho de que los moños serán iguales). Para elaborar su respuesta a las preguntas planteadas necesita, en primer lugar, "darse cuenta" de que debe dividir la cantidad de metros de listón por la cantidad de moños. En segundo lugar, debe realizar una división sencilla (en el primer caso, 12 dividido 8; en el segundo 2 dividido 3). Estas actividades han sido propuestas a alumnos de sexto grado de escuela primaria. En este nivel uno podría suponer que "darse cuenta" de que es necesario realizar una división no supone un desafío intelectual importante para el estudiante, especialmente porque se trata de un tipo de tarea ("ejercicio") que se repite con frecuencia en las aulas bajo diversos formatos, es decir, cambiando las situaciones y la información.

Importa especialmente señalar que estas dos actividades formaban parte de una prueba que contenía seis actividades. Todas ellas planteaban la misma situación y la misma pregunta, cambiando únicamente la cantidad de metros de cinta y la cantidad de moños. Por tanto, el foco de la propuesta no estaba puesto en la capacidad para "darse cuenta" de qué operación es necesario utilizar para resolver una situación, sino en la realización de sucesivas divisiones con números enteros de una o dos cifras. Lo que varía entre las distintas actividades es la dificultad de la división a realizar. En resumen, esta actividad evalúa la capacidad del estudiante para realizar divisiones sencillas.

Veamos ahora un segundo ejemplo (Figura 2.3). También en este caso la actividad forma parte de una prueba de evaluación para estudiantes de 6º grado de primaria. Esta actividad incluye una *situación*, variada *información* y varias *consignas*.

Figura 2.3: Actividad de evaluación en Matemática en 6º de Primaria



EVALUACIÓN

8/8

Lee con atención antes de resolver o contestar.
Recuerda seleccionar los datos pertinentes y realizar sólo lo que te piden.

1.- Salida de fin de año.
Precio por día y por niño en el complejo turístico: $ 320.
Costo total del transporte: $ 12810.
Precio de la entrada al parque acuático: $ 25 cada uno
Número de alumnos que participarán en la salida: 42.
Duración de la estadía: 3 días.
Averigua el costo total por alumno y lo gastado por todos los alumnos en los tres días de paseo
Indica con una frase lo que se busca con cada cálculo realizado.
Resuélvelo.

320 x 3 x 42 = A El precio total de lo que gastaron en el complejo todos los niños juntos.
25 x 42 = B El precio de todas las entradas juntas
A + B + 12810 = C Lo gastado por todos los alumnos juntos en los tres días
C : 42 = D Lo gastado por cada niño en los tres días

5/5

Fuente: Base de registros del estudio **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas**.

El planteamiento de la actividad comienza con tres indicaciones generales a los estudiantes: leer con atención, seleccionar los datos adecuados y enfocarse en realizar lo que se solicita en las consignas. Estas indicaciones son sumamente pertinentes, porque están dirigidas a prevenir algunas causas comunes de mal desempeño de los estudiantes en una prueba: leer demasiado rápido sin poner atención en qué se les pide que hagan, pasar rápidamente a "hacer operaciones" con todos los números que encuentran, sin detenerse a analizar cuidadosamente qué información es relevante para ello. Estos modos de actuar de los estudiantes muchas veces son propiciados por los docentes, cuando proponemos actividades como las de la Figura 2.2, que son reiterativas, que solo incluyen los datos estrictamente necesarios para resolverlas y que se resuelven mediante una única operación. Con ese tipo de actividades los estudiantes se acostumbran a proceder mecánicamente a realizar la operación que intuyen como más apropiada, con los datos que encuentran en la propuesta.

Luego de las indicaciones se pone a los estudiantes ante una **situación** propia de la vida real en la escuela: la realización de un paseo de fin de año. La situación es más plausible y más compleja que la planteada en la Figura 2.2 (hacer moños de listón). La cantidad de **información** que se da a los estudiantes es más amplia que en la actividad de la Figura 2.2. Se les indica el precio del alojamiento diario de cada niño en un complejo turístico, el costo

del transporte, el precio de la entrada a un parque acuático, la cantidad de alumnos que participarán en el paseo y la duración de la estadía.

A continuación se plantea a los estudiantes tres consignas: "*averigua el costo total por alumno y lo gastado por todos los alumnos en los tres días de paseo*"; "*indica con una frase lo que se busca con cada cálculo*" y "*resuelve los cálculos*". Sigue a continuación el planteamiento de cuatro operaciones aritméticas que involucran algunos de los datos incluidos como información. Lo que hace especialmente interesante a esta actividad es que la maestra no pide a los estudiantes que resuelvan directamente dichas operaciones, sino que primero les pide que indiquen "*qué se busca*" con cada una o, en otras palabras, para qué sirve realizarla, qué información nueva aportará realizar la operación en cuestión. Esto implica un requerimiento de reflexión por parte de los estudiantes sobre el sentido de las operaciones en la situación en cuestión. No se trata simplemente de "descubrir" qué operación tengo que realizar con los datos, sino de analizar qué pregunta puedo responder con cada operación.

La tarea tiene una complejidad adicional en cuanto al nivel de abstracción que deben manejar los estudiantes, derivada del uso de símbolos matemáticos que representan la nueva información derivada de cada cálculo. La primera operación tiene como resultado "A" y la segunda tiene como resultado "B". Luego "A" y "B" forman parte de la tercera operación, para dar como resultado "C", que a su vez es incluido en la cuarta operación.

Resulta bastante obvio que realizar la actividad propuesta en la Figura 2.3 requiere del estudiante un nivel de reflexión superior al que requiere resolver lo solicitado en la Figura 2.2. El estudiante tiene que manejar más información (cinco datos en la Figura 2.3, dos datos en la Figura 2.2); tiene que trabajar con tres operaciones (suma, multiplicación y división); tiene que inferir qué dato o información nueva surge de realizar cada operación indicada; tiene que analizar el significado de nuevos datos representados con letras (A, B, C); tiene que expresar su respuesta con sus propias palabras en una frase escrita. Las actividades incluidas en la Figura 2.2, en cambio, son relativamente más sencillas, apelando básicamente a la resolución de divisiones simples y a la expresión de la respuesta a través de un número. En resumen, las actividades de evaluación de las Figuras 2.2 y 2.3 difieren sustancialmente en lo que en adelante denominaremos "*requerimiento cognitivo*", es decir, el grado de complejidad de los procesos de pensamiento (reflexiones y razonamientos) que el estudiante debe realizar para resolverlas. Volveremos sobre el significado de la expresión "*requerimiento cognitivo*" un poco más adelante.

2. Caracterización de las actividades de evaluación según el formato de respuesta requerido

Decíamos que toda actividad de evaluación incluye una consigna o enunciado que le solicita al estudiante una respuesta, un resultado, una producción o un desempeño. Un primer aspecto a observar al analizar una actividad de evaluación es el tipo de respuesta que se pide al estudiante. La primera gran distinción que puede hacerse es entre tareas de **respuesta cerrada** y tareas de **respuesta construida**. En las primeras el estudiante debe seleccionar, señalar o expresar una respuesta que es única y predefinida; en las segundas el estudiante tiene que elaborar la respuesta.

Dentro de las tareas de **respuesta cerrada** el formato más típico y conocido son las preguntas de *múltiple opción*, en las que el estudiante debe seleccionar y marcar la respuesta correcta entre varias (Figura 2.4.A). Otras formas de tareas de respuesta cerrada son las de *apareamiento* (Figura 2.4.B.) y aquellas en que el estudiante debe indicar si una afirmación es *verdadera o falsa*. Consideraremos también como tareas de respuesta cerrada a aquellas que se responden con una palabra *predeterminada* como, por ejemplo, completar blancos o nombrar elementos en una figura o esquema (ver Figura 2.4.C.). Finalmente, incluimos en esta categoría[4] a los ejercicios que tienen un único resultado correcto y predefinido, como los incluidos en la Figura 2.4.D.

Figura 2.4: Ejemplos de tareas de respuesta cerrada

A: Selección múltiple
1 - ¿Qué postula Lamarck en su teoría sobre la evolución? a) El ambiente selecciona a los organismos más fuertes. b) Los organismos no cambian a través del tiempo. c) La selección natural es el mecanismo de la evolución. d) Las características adquiridas por uso y desuso de órganos se heredan. e) Las mutaciones son la causa de variabilidad genética.

4. En algunas clasificaciones este tipo de preguntas y el anterior se consideran como de respuesta abierta breve, porque el estudiante debe escribir la respuesta y no seleccionarla. Desde nuestra perspectiva, preferimos incluirlas entre las cerradas, porque no hay elaboración de la respuesta por parte del estudiante, dado que la misma está predeterminada.

B: Apareamiento

Evaluación escrita de Biología

Nombre	Grupo	Fecha

1. Busca la correspondencia entre el concepto y su significado y únelos con flechas:

1) EXTEROCEPTORES	A) Células encargadas de sostener, nutrir y proteger a las neuronas.
2) NEURONAS	B) Sustancias químicas encargadas de transmitir información de una neurona a otra.
3) NEUROTRANSMISORES	C) Células que reciben, conducen y transmiten el impulso nervioso.
4) NEUROGLÍAS	D) Estructuras nerviosas especializadas en captar estímulos provenientes del exterior.

C: Identificar elementos

5. Coloca la letra que corresponda de la imagen en cada nombre de las estructuras de la piel:
Epidermis____
Receptores sensoriales____
Hipodermis____
Glándula sudoripara____
Glándula sebácea_____
Dermis____
Pelo____

D: Ejercicios con resultado único predefinido

VI Serie
Instrucciones: Resuelva los siguientes problemas
1. Un niño camina $\frac{2}{3}$ Km de su casa al mercado y $\frac{1}{12}$ Km del mercado hacia el campo o fútbol. ¿Cuántos Km recorre en total?
2. Una doctora utiliza $2\frac{1}{2}$ litro de suero para un paciente y $1\frac{2}{6}$ litros para otro. ¿Cuántos litros de suero utiliza?
3. En una hora, Vicky corre $10\frac{1}{6}$ Km y Catalina $11\frac{3}{4}$ km. ¿Cuántos Km más corrió Catalina?
4. Una ingeniera civil dirige la construcción de una carretera que mide $45\frac{1}{2}$ de distancia. Si la obra ya lleva $24\frac{2}{5}$ Km, ¿Cuántos Km faltan?

Fuente: Elaboración propia a partir de las bases de registros de los estudios **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas** y **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina.**

Las tareas de **respuesta construida** incluyen también una amplia gama de posibilidades. Las más comunes son preguntas ante las que el estudiante debe redactar una respuesta breve cuyo contenido está predefinido, como las incluidas en la Figura 5.A. En otros casos la respuesta que debe producir el estudiante es más extensa, tal como ocurre en pruebas en las que el estudiante debe escribir un ensayo o similar como, por ejemplo, la consigna incluida en la Figura 5.B. Este formato "escrito", típico de la educación secundaria, requiere que el estudiante escriba todo lo que "sabe" sobre un tema. Suele utilizarse en disciplinas como Filosofía, Derecho y Literatura.

En este punto es importante explicar la diferencia entre tareas de respuesta **construida** y tareas de respuesta **abierta**. Muchas veces se utiliza ambos términos de manera intercambiable. Sin embargo, la expresión "abierta" tiene un doble significado.

1. Por un lado, una tarea es de respuesta "abierta" en la medida en que el estudiante debe elaborar y redactar la respuesta. En este sentido "respuesta abierta" y "respuesta construida" son equivalentes. Desde este significado, las preguntas de respuesta "construida" pueden diferenciarse en función de su extensión. Algunas requieren una elaboración breve (Figura 5.A), otras una elaboración más extensa (Figura 5.B).
2. Por otro lado, una tarea es de respuesta "abierta" en la medida en que su contenido no está predefinido. El estudiante debe crear o producir algo nuevo. Es "abierta" en tanto admite diversidad de respuestas. No hay una única respuesta apropiada, sino que existe un espacio más o menos importante para la toma de decisiones y la creación por parte del estudiante. Un ejemplo es el incluido en la Figura 5.C. En este caso, si bien no se dice explícitamente, la respuesta requerida es más bien breve. En otros casos la respuesta "abierta" puede ser extensa como, por ejemplo, cuando se solicita la redacción de un artículo (véase el ejemplo incluido en la Figura 5.D, que se presenta en forma completa más adelante en la Figura 2.16) o de un proyecto de investigación o similar.

Figura 2.5: Ejemplos de tareas de respuesta abierta

5.A: Respuesta construida predefinida breve	5.B: Respuesta construida predefinida extensa
PREGUNTAS 1. ¿Qué son las hormonas? 2. ¿Cómo podemos controlar las hormonas? 3. ¿En qué nos afectan las hormonas? 4. ¿Cuáles son las principales hormonas? 5. ¿En qué parte del cuerpo encontramos las hormonas?	***"Explique el funcionamiento de los poderes ejecutivo, legislativo y judicial, y las relaciones entre ellos"***
5.C: Respuesta construida abierta breve **2. Analiza el siguiente texto:** *"En un área verde, con gran variedad de especímenes de flora y fauna autóctonas, e ha previsto la construcción de un Shopping.* *El proyecto implica, por un lado, la destrucción de las tres cuartas partes de la flora del lugar y del hábitat de las especies animales, pero, por otro lado es una fuente importante de mano de obra porque se contrataría gran cantidad de obreros del ámbito local, que, dado el reciente cierre de fábricas han quedado sin trabajo".* a. ¿Cuál es el problema que se plantea en el texto? b. Desde la perspectiva del desarrollo sustentable, ¿qué posibles soluciones darías al problema? Explica.	
5.D: Respuesta construida abierta extensa Suponiendo que tú fueras un periodista que escribe una columna de divulgación científica en un semanario, deberás escribir un artículo que se titule: "Calentamiento global y desarrollo industrial: ¿juntos o separados?" Dicho artículo deberá tener las siguientes características: • Contener argumentos a favor y/o en contra de esta relación • Incluir la gráfica que se presenta a continuación y hacer referencia expresa a ella en el texto. • Hacer mención de la Revolución Industrial como hecho histórico en relación con la variación de la concentración de dióxido de carbono en la atmósfera.	

Fuente: Elaboración propia a partir de las bases de registros de los estudios **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas** y **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina**.

Las actividades de respuesta construida se definen por el hecho de que el estudiante debe elaborar la respuesta y varían en función de dos variables principales: la extensión de la respuesta (breve, intermedia, extensa) y la apertura del abanico de respuestas posibles (el grado en que el estudiante debe crear buena parte del contenido de su respuesta y no simplemente reproducirlo a partir de lo que ha estudiado antes). Obviamente las fronteras entre estas categorías son borrosas, pero es útil contar con una clasificación. La Tabla 2.1 ofrece una sistematización de estos distintos formatos de respuesta y constituye un primer "lente" para mirar las actividades de evaluación, tanto para revisar las que estamos utilizando, como para tenerlo en cuenta en la elaboración de nuevas actividades. La pregunta a hacerse es qué tipo de respuesta espero de los estudiantes y por qué ese tipo es adecuado a la finalidad de mi evaluación.

Tabla 2.1: Categorización de tareas de evaluación según formatos de respuesta requeridos

Respuesta cerrada	Respuesta construida
Marcar la respuesta:	**El contenido de la respuesta está predefinido:**
• Múltiple opción	• Extensión breve (algunas frases)
• Verdadero-Falso	• Extensión media (algunos párrafos)
• Apareamiento de columnas	• Extensión amplia (varias páginas)
Escribir la respuesta:	**El contenido de la respuesta está abierto:**
• Completar blancos	• Extensión breve (algunas frases)
• Escribir una palabra o frase	• Extensión media (algunos párrafos)
• Identificar y nombrar elementos	• Extensión amplia (varias páginas)
• Dar un resultado numérico	

Fuente: Elaboración propia.

Es importante notar que no existe correlación estricta entre el formato de la respuesta y la complejidad o nivel de reflexión que exige la tarea. Normalmente se piensa que las preguntas de respuesta cerrada son memorísticas y mecánicas, y que no apelan a la reflexión, en tanto que las de respuesta construida sí lo hacen. Un poco más adelante en este capítulo analizaremos esta cuestión. Veremos como ciertas preguntas de respuesta cerrada bien elaboradas pueden requerir comprensión y reflexión por parte del estudiante, al tiempo que muchas tareas de respuesta construida se pueden responder repitiendo de memoria conceptos o explicaciones estudiadas en clase.

Las tareas de respuesta cerrada tienen como principal ventaja para el docente el hecho de que pueden ser corregidas en forma rápida, a diferencia de las de respuesta abierta. Sin embargo, la contracara de esta ventaja es que construir buenas tareas de formato cerrado requiere tiempo y análisis cuidadoso. Las tareas de respuesta construida tienen como principal ventaja el hecho de que requieren que el estudiante "elabore" una producción. No tiene la posibilidad de responder por azar o por descarte. Asimismo, su elaboración requiere, en principio, menos tiempo que las tareas de respuesta cerrada. Como contrapartida, si están bien elaboradas, su corrección requiere más tiempo porque habrá diversidad de producciones y respuestas, de mayor extensión y más difíciles de valorar. De modo que las tareas de respuesta abierta requerirán de mayor inversión de tiempo en la elaboración de pautas

de valoración y en la revisión de las mismas[5]. No es fácil corregirlas de un modo claro y transparente para los estudiantes si no hay pautas o rúbricas "explícitas", que vayan más allá de establecer "cuántos puntos vale" la respuesta a cada pregunta.

Es bastante común que las consignas o enunciados tengan debilidades en su formulación, a las que es preciso prestar atención. Las actividades de respuesta cerrada suelen tener como debilidad principal el hecho de que incluyen alternativas que, por el modo en que están formuladas, son fácilmente descartables utilizando el sentido común. A veces, incluso, por aspectos de concordancia en el lenguaje, entre la pregunta y las opciones de respuesta. En el caso de las actividades de respuesta abierta, las principales debilidades tienen relación con su falta de claridad acerca de qué es lo que se espera del estudiante y/o insuficiente explicitación de las características de la producción o respuesta esperada. Muchas veces los docentes damos por sentado, sin explicitarlo, muchos aspectos de lo que esperamos de los estudiantes al realizar la actividad.

3. Información y situaciones involucradas en las actividades de evaluación

En el apartado anterior nos enfocamos en el análisis de las respuestas o producciones que las consignas de evaluación solicitan a los estudiantes. En este apartado nos enfocaremos en el segundo aspecto clave de las actividades de evaluación: el tipo de información y situaciones que ponen ante los estudiantes y que estos deben procesar.

Según se señaló más arriba, buena parte de las actividades de evaluación incluyen información que el estudiante debe leer, analizar y comprender. Si bien en algunos casos se proponen preguntas sin más información que la propia pregunta (Figuras 2.4.A, 2.4.B, 2.5.A y 2.5.B), en otros casos las actividades incluyen información que el estudiante debe analizar para construir o elegir su respuesta (Figuras 2.2, 2.3, 2.4.C, 2.4.D, 2.5.C y 2.5.D). La información puede estar presentada como un texto en prosa o bajo otros formatos: datos numéricos, tablas, gráficos, esquemas y/o ilustraciones, por mencionar algunos. En general la presencia de información implica un desafío mayor para el estudiante, ya que deberá procesarla y comprenderla. Al mismo tiempo, un exceso de información puede ser contraproducente y confundir al estudiante como, por ejemplo, textos muy extensos o gráficos

5. El tema de las rúbricas y su construcción y uso, tanto con fines formativos como de calificación, se analizará en los capítulos 4 y 5.

con mucha información (dependiendo, obviamente, del nivel educativo de que se trate).

Algunos docentes argumentan que si se emplea información, en particular textos, en evaluaciones que no son de lenguaje sino de otras disciplinas (por ejemplo, matemática o ciencias), en realidad no se está evaluando el contenido de la disciplina sino la capacidad de lectura. Este argumento es altamente discutible y depende, en parte, de cómo uno defina el aprendizaje en la disciplina en cuestión. Si por hacer matemática se entiende resolver operaciones y aplicar propiedades, probablemente la presencia de texto "distraiga" o complejice innecesariamente la tarea del estudiante. Pero, si por hacer matemática se entiende la capacidad de modelizar situaciones y traducirlas a términos matemáticos, entonces la capacidad de lectura para comprender la situación es parte esencial del aprendizaje de la disciplina.

Desde nuestra perspectiva, en todas las disciplinas el aprendizaje requiere de un nivel de lectura aceptable. Comprender información sobre la disciplina en distintos formatos es parte de la formación en la misma. La lectura es una habilidad transversal a todas las disciplinas y debería ser trabajada en todas las asignaturas, no es algo "dado" o adquirido de una vez y para siempre por los estudiantes. Lo mismo ocurre con la escritura. Cada disciplina tiene sus propios estilos, tipos de textos y formas específicas de lectura y escritura. De cualquier manera, es importante elegir cuidadosamente los textos y la información a incluir en las tareas, para que no resulten confusos, excesivamente extensos o poco motivadores.

La información que se presenta a los estudiantes en las tareas de evaluación implica siempre algún tipo de **situación**, entendida como conjunto de factores o circunstancias que contextualizan la labor del estudiante. La primera "situación" obvia en que se encuentra el estudiante es la propia evaluación: el estudiante debe realizar una tarea que le propone su docente con el fin de valorar lo que ha aprendido para calificarlo y/o ayudarle en su proceso de aprendizaje. Esta es la situación básica de evaluación, que tiene sus matices según la finalidad de la evaluación y el estilo del docente.

Hecha esta salvedad, es importante notar que toda actividad de evaluación que incluya información a procesar, es decir, que no se limite a plantear una pregunta, implica la existencia de un conjunto de elementos que constituyen una situación, dado que ponen al estudiante en un contexto. Este contexto o situación puede ser de distinta naturaleza: propio de la vida cotidiana inmediata, de la vida en sociedad o de la actividad propia de la disciplina estudiada.

A los efectos del análisis de las tareas de evaluación es relevante identificar cuatro principales tipos de situaciones o contextos (Ravela y otros 2014).

1. Situaciones de la vida cotidiana[6]: corresponden a las actividades que llevan adelante los estudiantes, vinculadas con los deportes, la música, la vida en el centro educativo, los amigos, en entretenimiento o la familia por mencionar los principales. La actividad presentada en la Figura 2.3 es un ejemplo de situación propia de la vida cotidiana en la escuela: la organización de un paseo grupal.
2. Situaciones sociales: si bien son también parte de la vida cotidiana, están vinculadas con problemas de mayor envergadura, que involucran a toda la sociedad, como las cuestiones ambientales, las económicas, las vinculadas con la pobreza, el uso de la energía o tantas otras situaciones relacionadas con la vida social. La actividad presentada en la Figura 2.5.C. es un ejemplo de este tipo de situaciones.
3. Situaciones disciplinares: son aquellas que reflejan problemas y prácticas posibles en la producción y uso del conocimiento en una disciplina. Colocan al estudiante ante algunas de las formas de actividad y problemas propios de quienes trabajan realmente en la producción de conocimiento en la disciplina o en el uso del mismo para resolver problemas reales. Dentro de esta categoría se incluyen situaciones como la demostración de un teorema matemático; situaciones simuladas de investigación que ponen al estudiante ante un experimento narrado y que requieren un razonamiento o diseño de tipo experimental; o propuestas que incluyen la observación y registro de datos de la realidad, y/o la manipulación de materiales reales (como, por ejemplo, la disección de un encéfalo de res).
4. Situaciones escolares: se trata de situaciones que son creadas por los docentes con la intención de dar un contexto a la actividad de evaluación, pero que en realidad no reflejan adecuadamente el uso del conocimiento en la vida real ni en la actividad propia de la disciplina (si bien incluyen elementos de ello). Son intentos por acercarse a la realidad, que denominaremos "escolarizados". En algunos casos se trata de situaciones triviales[7] que intentan utilizar elementos de la vida cotidiana (como, por ejemplo, "hacer moños" en la Figura 2.2 o las situaciones propuestas en los ejercicios de la Figura 2.4.D). En otros casos se trata de intentos de proponer

6. Cotidiano en el sentido de propio de la vida ordinaria, no en el sentido de que tenga lugar todos los días.
7. De acuerdo con la RAE trivial es algo sabido de todos, común y que carece de toda importancia y novedad.

situaciones disciplinares, pero que no reflejan usos reales del conocimiento en las mismas.

La cuestión de las "situaciones" que se proponen es de fundamental importancia. A este tema dedicaremos por entero el capítulo que sigue. En este momento nos interesa explicar la distinción principal entre lo que denominaremos "situaciones auténticas" y "situaciones escolares". En ambos casos se puede tratar de situaciones referidas a contextos de la vida cotidiana, a asuntos de la vida social o a los temas y prácticas en una disciplina. La cuestión fundamental que distingue a unas de otras es el grado en que el conocimiento o los procedimientos que forman parte de la actividad realmente se ponen en juego y se utilizan de esa manera en la vida real (personal, social o disciplinar). Para explicar mejor la idea recurriremos a dos pares de ejemplos.

En primer término veamos dos ejemplos de situaciones referidas a la vida cotidiana (Figuras 2.6 y 2.7). En ambos casos se trata de actividades que remiten al estudiante a una situación de la vida cotidiana: repartir chocolate. La finalidad es que los estudiantes realicen ciertos razonamientos y operaciones aritméticas. La diferencia fundamental radica en que la situación planteada en la Figura 2.6 difícilmente se verifique en la vida real. Es una situación ficticia. Nadie "mide" el chocolate que consume en fracciones ni se plantea las preguntas y operaciones que esta actividad propone. En cambio, la actividad propuesta en la Figura 2.7 sí refleja una situación de la vida cotidiana en que la que puede ser necesario emplear los razonamientos y cálculos indicados.

Figura 2.6: Actividad "escolar" de reparto de chocolates en Matemática, 6º de primaria

Leé atentamente y resolvé

1- Laura comió $\frac{5}{9}$ de un chocolate y le convidó a su hermano **la mitad** de lo que le quedaba.

a- ¿Qué fracción del chocolate comió el hermano de Laura?

b- ¿Qué fracción del chocolate comieron entre los dos?

Fuente: Base de registros del estudio **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas.**

Figura 2.7: Actividad "auténtica" de reparto de chocolates en Matemática, 6º de primaria

> Para repartir 23 chocolates entre 5 niños, Vanesa pensó lo siguiente:
> "23 chocolates entre 5 me da 4 chocolates para cada uno, pues 4x5 = 20 y me sobran 3 chocolates que los corto cada uno en cinco partes y entrego una parte de cada chocolate a cada uno".
>
> En cambio Joaquín lo pensó así:
> "Le doy 4 chocolates a cada uno igual que Vanesa pero con los 3 chocolates que quedan, corto cada uno por la mitad y le doy una mitad a cada niño, luego divido el último medio en cinco y le doy una parte a cada uno".
>
> Analiza si son o no equivalentes los repartos de Vanesa y Joaquín.

Fuente: Base de registros del estudio **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas**.

Algo parecido ocurre con las actividades incluidas en las (Figuras 2.8 y 2.9). En este caso de trata de situaciones propias de disciplinas científicas, la física y la biología.

Figura 2.8: Actividad "escolar" con contenidos de física

> 2) Un gato se mueve persiguiendo a un ratón sobre una trayectoria rectilínea.
> El gato comienza a perseguir al ratón desde el reposo y alcanza una velocidad de 8,0 m/s en 5,0 s.
> Luego continúa moviéndose con velocidad constante durante 4,0 s más.
> A partir de ese instante comienza a moverse con una aceleración de -0,5 m/s/s, hasta detenerse.
> En estas condiciones determina:
> a) La aceleración del gato en los primeros 5,0 s
> b) La velocidad del gato a los 12 s
> c) ¿A qué distancia del punto de partida se detuvo el gato?
> d) Si el ratón se encontraba a 3,0 m del gato cuando lo comenzó a perseguir. ¿Con qué velocidad constante mínima debió correr para que el gato no lo atrapara en este movimiento?

Fuente: Base de registros del estudio **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas**.

En el primer caso (Figura 2.8), se trata de la aplicación de conceptos de física (velocidad, distancia, tiempo y aceleración) a una situación cotidiana (un gato que corre persiguiendo a un ratón). Consideramos a esta situación como "escolar", porque difícilmente alguien utilice los conocimientos involucrados, propios de la física clásica, para analizar una situación como esta.

Se trata de un ejercicio en el que el estudiante debe aplicar las fórmulas y conceptos estudiados en clase, pero no de una situación auténtica. Los científicos utilizan estas fórmulas, en todo caso, para analizar el movimiento de cuerpos celestes, en balística o en otro tipo de situaciones.

La actividad incluida en la Figura 2.9, en cambio, propone una situación "auténtica" relacionada con la genética, en el sentido de puede ocurrir en la vida real y que los conceptos y leyes implicados son relevantes para resolver la situación planteada. Es propiamente un "problema", en el sentido de que colocan a los estudiantes en una situación nueva, relevante y posible, en la que es necesario y pertinente recurrir a los contenidos estudiados en clase.

Figura 2.9: Actividad "auténtica" con contenidos de Genética

11. Unos investigadores colombianos han tratado de encontrar las bases genéticas de la coloración de las alas de dos especies de mariposas, muy abundantes en el Valle del Cauca. La siguiente tabla muestra algunos alelos responsables de la coloración de las alas en estas mariposas:

ALELO	CARACTERÍSTICAS
A	ala negra con una mancha blanca
a	ausencia de manchas blancas en el ala negra
B	línea gruesa amarilla debajo de las manchas blancas
b	ausencia de línea gruesa amarilla debajo de las manchas blancas
C	ausencia de punto rojo sobre las manchas blancas
c	presencia de punto rojo sobre las manchas blancas

Los investigadores notaron que las mariposas Sp1 cuya coloración alar era completamente negra, tenían una mayor probabilidad de sobrevivir al ataque de los predadores pues eran más difíciles de detectar dentro del bosque en el que habitan. Así mismo, encontraron que algunos individuos de Sp2 presentaban una coloración similar a la de Sp1 para reducir el riesgo de ser atacadas por los predadores. Si esto es así, usted podría proponer que los genes que producen la coloración de estas mariposas imitadoras (Sp2) son

A. AA BB CC **B. aa bb CC** **C. AA bb cc** **D. aa bb cc**

Fuente: Base de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina**.

En este punto es relevante realizar una distinción entre "ejercicio" y "problema". Muchos docentes utilizan estos dos términos como sinónimos. Sin embargo, refieren a tipos de actividad completamente diferentes. El término "ejercicio" remite a la aplicación de conceptos o procedimientos conocidos en situaciones conocidas y reiteradas. El término "problema", en cambio, remite a la resolución de situaciones nuevas, que normalmente no tienen una única forma de solución. *"«Ejercicio» es toda propuesta que apela a la aplicación de procedimientos conocidos por el sujeto que debe resolverlos. «Problema» es toda situación que involucra conocimientos de uno o más campos, que plantea un interrogante, que admite más de una perspectiva de análisis y que*

no ha sido resuelta anteriormente por el sujeto" (Picaroni y Loureiro 2010: 38). Refiriéndose a la enseñanza de la matemática, agregan:

> *"La distinción entre ejercicio y problema no es absoluta ya que una misma situación matemática puede entenderse como lo primero o como lo segundo, en función de a quién se destina... No existen pues, ejercicios y problemas en forma independiente de las personas que deben resolverlos. A la pregunta «¿cuántos lápices le corresponden a cada uno de los 16 alumnos del grupo si reparto en partes iguales los 240 de una caja?», algunos responderán «15» en forma inmediata aplicando la división 240:16 exitosamente, y otros tardarán cierto tiempo pues deberán buscar entre los conocimientos que disponen una estrategia adecuada para resolver algo a lo que nunca se han visto enfrentados".*

La importancia y el proceso de elaboración de problemas y situaciones auténticas será objeto de trabajo en el capítulo que sigue.

Lo trabajado a lo largo de este apartado permite contar con un segundo "lente" para analizar nuestras propuestas de evaluación, en este caso desde la perspectiva de la información que los estudiantes deben procesar y, principalmente, enfocándose en los tipos de situaciones que les proponemos. Las preguntas principales a hacerse son:

1. La información que debe ser analizada para realizar la actividad, ¿es clara, motivadora, suficientemente desafiante pero, al mismo tiempo, no excesivamente compleja?
2. ¿A qué ámbito de actividad corresponde la situación (cotidiana, social, disciplinar o de otro tipo)?
3. ¿Se trata de situaciones nuevas que implican un problema a resolver o de situaciones conocidas que requieren repetir o "ejercitar" procedimientos?
4. ¿Se trata de situaciones "auténticas", en el sentido de que es razonable utilizar en la vida real los conceptos o procedimientos involucrados, de la manera requerida y ante una situación parecida?

Para que el lector tenga una visión de cuáles son las prácticas predominantes en la región, sobre un total de 4359 tareas de ciencias naturales de 3º de secundaria básica relevadas en Cali, Lima, Montevideo y Santiago de Chile en 2012, solo el 37% incluía algún tipo de información o situación a procesar. Esto es, el 73% eran preguntas directas, sin información adicional para leer y analizar. Asimismo, el 81% carecía de una situación o contexto. Del restante 19% que sí ponía a las estudiantes en una situación, la distribución fue la siguiente: 10% eran situaciones escolares, 4% eran situaciones

auténticas cotidianas, 3% eran situaciones auténticas propias de la disciplina y 2% eran situación sociales auténticas (Ravela y otros, 2014).

4. Aprendizaje superficial y aprendizaje profundo

En los apartados anteriores analizamos las actividades de evaluación a partir de la caracterización de cada uno de sus dos componentes principales: el tipo de respuesta o producción que requieren del estudiante, por un lado, y los tipos de información y situaciones que deben procesar, por otro. Cada vez que proponemos una actividad a nuestros estudiantes, sea como parte de una evaluación o, sencillamente, como parte de nuestra propuesta de enseñanza, estamos colocándolos en una situación en la que deben "procesar" algún tipo de información o analizar una situación, con el fin de elaborar una respuesta.

En este apartado y el que sigue nos enfocaremos en un análisis más abarcador de los "procesos cognitivos" involucrados en la realización de las actividades de evaluación.

Los seres humanos tenemos la capacidad de percibir las realidades a nuestro alrededor -realidades materiales y sociales- y en nuestra propia persona -realidades físicas y emocionales-, capacidad que compartimos con el resto de los animales. Pero los seres humanos podemos además construir representaciones conscientes de dichas realidades. Para ello nuestro cerebro tiene la capacidad, a través del lenguaje, de nombrar las percepciones, establecer relaciones entre ellas y construir imágenes complejas del mundo material, social y personal. A estos procesos de construcción intelectual que nos permiten describir, explicar y valorar las realidades naturales, personales y sociales, los denominamos en este libro "procesos cognitivos".

La expresión *cognitivo* fue utilizada por Jean Piaget (1973) en la primera mitad del siglo XX, quien se enfocó en mostrar cómo las capacidades intelectuales se desarrollan a lo largo de la vida de los seres humanos, partiendo de las habilidades motrices y de la manipulación de objetos reales en los primeros años de vida para, progresivamente, ir constituyéndose en operaciones lógicas de abstracción creciente. Piaget enfatizó una visión del aprendizaje como construcción de significados por parte del individuo, por oposición a otras en las que el aprendizaje es predominantemente una cuestión de automatización de respuestas ante los estímulos del ambiente, o de adquisición y memorización de contenidos. Para estas últimas, el conocimiento está dado, preestablecido, y el estudiante debe "adquirirlo" y demostrar que lo ha hecho a través de la evocación y la enunciación del mismo.

Posteriormente David Ausubel (1978) desarrolló el concepto de *aprendizaje significativo,* mostrando la importancia de que el contenido a aprender se "inserte" de alguna manera en la trama de significados e ideas que el estudiante trae consigo. El aprendizaje se produce cuando los temas tienen "sentido" para el estudiante, cuando este puede comprender su significado. Para comprender un concepto nuevo es necesario que el estudiante pueda relacionarlo con sus ideas previas e integrarlo a su propia construcción intelectual sobre el mundo.

En los años 70 Ference Marton y Roger Säljö (1976) desarrollaron la distinción entre *aprendizaje superficial* y *aprendizaje en profundidad*[8]. Partieron de un estudio cualitativo acerca de los procesos de pensamiento de estudiantes universitarios en torno a una actividad de evaluación: los estudiantes debían leer un texto razonablemente sencillo sobre una propuesta de reforma universitaria y responder preguntas de comprensión. Lo que llamó la atención de los investigadores fue la diversidad de interpretaciones de los estudiantes, algunas claramente opuestas a la intención del autor del texto. El análisis de la manera en que cada estudiante procesó el texto los condujo a una conclusión obvia y sorprendente a la vez: "*los estudiantes que no llegaron a una respuesta apropiada simplemente no la estaban buscando*" (Marton y Säljö, 2005).

> "*La principal diferencia que encontramos en el proceso de aprendizaje fue si los estudiantes se enfocaron en el texto en sí mismo o en sobre qué era el texto (la intención del autor, el punto principal, la conclusión a extraer). En el primer caso el foco de atención eran las páginas del texto, y en el segundo estaba más allá de las mismas. El primer modo de aproximarse a la tarea de aprendizaje se caracterizó por ser un esfuerzo ciego y espasmódico por memorizar el texto; estos estudiantes parecían, hablando metafóricamente, verse a sí mismos como vasijas vacías a ser llenadas con las palabras del texto. En el segundo caso, los estudiantes intentaron comprender el mensaje buscando relaciones al interior del texto o relaciones entre el texto y el mundo real, o buscando relaciones entre el texto y su estructura subyacente. Estos estudiantes parecían verse a sí mismos como creadores de conocimiento, que debían usar sus capacidades para construir juicios críticos, conclusiones lógicas y expresar sus propias ideas.* (Marton y Säljö, 2005: 43).

8. "*Las investigaciones de Ference Marton y Roger Säljö... fueron las primeras en diferenciar entre estudiantes que adoptan un enfoque profundo, cuando procuran comprender el significado, y otros que asumen un enfoque superficial, cuando solo se preocupan por memorizar el material*" (Camilloni y otros, 2010).

A estos dos modos de aproximarse al texto los denominaron respectivamente aprendizaje superficial y aprendizaje en profundidad. Esta distinción fue desarrollada posteriormente por distintos autores y trabajos.

El aprendizaje superficial, que se corresponde con lo que en este libro denominaremos actividades de "bajo requerimiento cognitivo", se enfoca en la reproducción de información. Los estudiantes se esfuerzan por memorizar la información que consideran necesaria para la instancia de evaluación, como un fin en sí mismo; intentan repetir lo que han "aprendido" tal como aparece en los textos, palabra por palabra; tienen una perspectiva estrecha y enfocada en detalles y no logran distinguir entre principios y ejemplos. Se limitan a recordar y repetir "lo que dice" fulano o la explicación dada por el profesor en el aula. Muchos estudiantes, desde la escuela primaria hasta la universidad -incluidos los posgrados- se acostumbran a memorizar datos, hechos, nombres de fenómenos, nombres de autores, definiciones y principios, y repetirlos sin comprender su significado. Subrayamos esta última expresión porque el problema no es el hecho de memorizar en sí mismo -la memoria es una de las capacidades principales de los seres humanos- sino la falta de significado, el hecho de recordar y repetir sin saber de qué se está hablando. A esta forma de aprendizaje se la suele denominar "conocimiento declarativo". Por lo general estos estudiantes son pasivos, tienen poca autonomía y poco interés por la materia, se limitan a estudiar lo estrictamente requerido en el curso y su principal motivación está vinculada con el miedo a fracasar o fallar en la evaluación (Lublin, 2003) (Cohen, Manion y Morrison, 2004).

El aprendizaje profundo, que se corresponde con lo que en este libro denominaremos como "alto requerimiento cognitivo" o "procesos cognitivos complejos", se caracteriza por buscar la comprensión del material estudiado; por una mirada abarcadora y amplia en torno a los temas que se estudian; y por el establecimiento de relaciones entre los temas nuevos y los conocimientos previos, así como entre los conceptos estudiados y situaciones o experiencias de la vida real. Por lo general los estudiantes con estilo de aprendizaje profundo están motivados por su propio interés en el tema o la materia, más que por la evaluación externa. Tienden a leer y estudiar más allá de lo estrictamente requerido para el curso (Lublin, 2003) (Cohen, Manion y Morrison, 2004).

Una línea de trabajo posterior, que está vinculada con la noción de aprendizaje "profundo", es el enfoque denominado "enseñanza para la comprensión", desarrollado por Howard Gardner y otros (Stone Wiske, 1999). Según este enfoque, en educación es más importante el tratamiento en profundidad de pocos temas, que el aprendizaje superficial y memorístico de

muchos[9]. El propósito principal de la educación debería ser que los estudiantes comprendan aquello que estudian y desarrollen en términos generales, una capacidad y actitud de comprensión. Wiggins (1998) define la comprensión a partir de cinco elementos centrales:

e. La capacidad de interpretar y explicar el material estudiado. Alguien comprende algo cuando es capaz de explicarlo a otro con sus propias palabras, de una manera original y clara, que va más allá de lo obvio, de la descripción superficial o de la reiteración más o menos literal de lo leído. Una buena explicación involucra ideas y "lecturas" personales del fenómeno explicado.
f. La capacidad de utilizar el conocimiento en diversidad de contextos, especialmente en contextos nuevos y demandantes, con el fin de resolver problemas y situaciones de un modo innovador, creativo y flexible.
g. La capacidad de adoptar múltiples perspectivas sobre la realidad. Alguien comprende en la medida en que tiene conciencia de que existen diversos puntos de vista sobre cada realidad y que no hay un saber o verdad única y acabada. Esto implica ser consciente de que el saber es siempre parcial y debe ser argumentado y sustentado más que enunciado o citado. Como consecuencia, la persona puede hacer y aceptar críticas, asume que todas las perspectivas son parciales y que puede haber distintos puntos de vista válidos sobre la realidad.
h. La capacidad de empatía. La comprensión implica ponerse en el lugar del otro y ver la realidad desde su perspectiva. Implica una salida del egocentrismo y la conciencia de que los demás experimentan y perciben la realidad de un modo diferente al propio. La persona tiene la capacidad de colocarse en la perspectiva de los demás y captar, al menos en parte, lo que otros sienten y piensan.
i. El autoconocimiento. Alguien comprende en la medida en que es consciente de sus propias limitaciones y puede darse cuenta de lo que sabe y lo que no. Como consecuencia, puede reconocer sus propios prejuicios. Por oposición, alguien que no comprende no tiene conciencia de sus propias limitaciones y prejuicios ni del rol que estos juegan en las propias opiniones.

---

9. Con relación a esta tensión entre extensión y profundidad en el tratamiento de los contenidos de la enseñanza, en el capítulo 3 veremos con más detalle los planteamientos de Gardner.

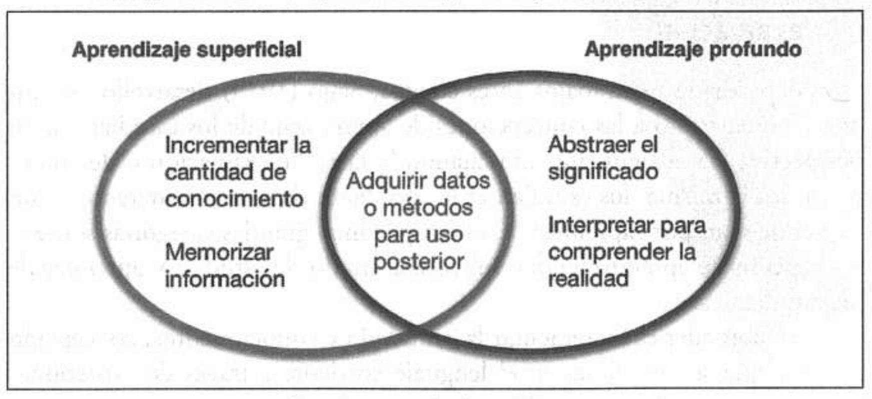
5. Los procesos cognitivos involucrados en las tareas de evaluación

A partir de sus trabajos antes citados, Säljö (1979) desarrolló una indagación en torno a las concepciones de aprendizaje de los estudiantes. Su perspectiva era la siguiente: aún cuando la tarea de evaluación o de aprendizaje sea la misma, los estudiantes la abordarán de manera diferente según qué entiendan por "aprender". Construyó cinco grandes categorías sobre la concepción de aprendizaje que subyace al modo de estudiar y aprender de los estudiantes:

- aprender es incrementar la cantidad de conocimientos, concepción que se manifiesta en el lenguaje corriente a través de expresiones como "saber mucho";
- aprender es memorizar, concepción que se materializa en la práctica de que un amigo o familiar ayude al estudiante "tomándole la lección" (el estudiante enuncia los contenidos y el "asistente" contrasta lo que dice con lo que está en el texto);
- aprender es adquirir datos o procedimientos para usarlos posteriormente, concepción vinculada con una visión del conocimiento como algo útil, que "sirve para" resolver problemas de distinto tipo;
- aprender es abstraer el significado, concepción que implica que el estudiante se da cuenta de que hay un significado por detrás de las palabras y textos que estudia y que estudiar implica desentrañar esos significados;
- aprender es un proceso interpretativo dirigido a comprender la realidad, concepción que implica que para estos estudiantes el conocimiento adquiere sentido en la medida en que les permite "entender" lo que ocurre en el mundo o lo que les ocurre a ellos mismos (si bien difícilmente el estudiante sea plenamente consciente como expresarlo en estos términos).

Un estudio posterior de Van Rossum y Schenk (1984) indagó la relación entre estas cinco concepciones de aprendizaje y los enfoques superficial o profundo que los estudiantes adoptan ante las tareas de evaluación. El estudio encontró una clara relación entre las dos primeras concepciones y el aprendizaje superficial, por un lado, y entre las dos últimas concepciones y el aprendizaje profundo por otro. La concepción del aprendizaje como adquisición de datos y métodos para uso posterior quedó en una situación intermedia y fue compartida por ambos "tipos de aprendices".

Figura 2.10: Concepciones de aprendizaje de los estudiantes



Fuente: Elaboración propia a partir de Van Rossum y Schenk (1984).

Estas concepciones del aprendizaje, que originalmente fueron estudiadas como "características" de los estudiantes, no son rasgos individuales innatos, que vengan con la persona, sino que son, en buena medida, resultado de las experiencias promovidas por el sistema educativo y por la cultura. El hecho de que un estudiante adopte uno u otro enfoque está determinado por sus experiencias escolares y por los mensajes que recibe de sus docentes y familiares. Por ejemplo, la práctica de "tomar la lección", que requiere estudiar de memoria, suele ser promovida a veces desde el ámbito escolar y, muchas otras, desde el hogar, como resultado de las experiencias escolares que en su momento tuvieron los padres. Muchas madres genuinamente preocupadas por apoyar a sus hijos destinan gran cantidad de tiempo a intentar "ayudarles" tomándoles la lección.

Una de las principales formas a través de las cuales la cultura escolar inculca en los estudiantes una u otra concepción del aprendizaje, son las actividades de evaluación y las prácticas que giran en torno a ellas, como la calificación y los exámenes. Es en este sentido que afirmamos que las actividades y consignas de evaluación son "reveladores" de qué es lo que está siendo priorizado como aprendizaje. De allí la importancia de su análisis, que es el propósito central de este capítulo[10].

10. Deliberadamente evitaremos ingresar en una discusión teórica en profundidad con respecto a una clasificación exhaustiva de estos procesos o sobre cuestiones en boga como "enseñanza por competencias", ya que muchas veces estas discusiones quedan muy alejadas de la realidad del aula y suelen terminar en cambios superficiales en el lenguaje utilizado para nombrar las cosas (y en modas relativas al uso de ciertos términos), más que en cambios en las prácticas de enseñanza.

Con el propósito indicado, utilizaremos como herramienta de análisis una categorización de las actividades de evaluación en función de los "procesos cognitivos" que deben ponerse en juego para resolverlas. La pregunta clave de este análisis es qué concepción de aprendizaje y qué tipo de relación con el conocimiento promueve cada tarea de evaluación. Para responder esta pregunta seguiremos, con algunas adaptaciones, la línea de trabajo que empleamos en investigaciones previas sobre el tema, en las que nos hemos enfocado en cinco grandes categorías de procesos cognitivos crecientemente complejos (Leymonié y otros, 2013). A continuación presentamos brevemente las cinco categorías, para luego ilustrarlas con actividades concretas recogidas en las aulas, en el marco de los estudios realizados en 2008 y 2012.

Actividades que requieren recordar y reproducir. Incluimos en esta categoría las preguntas cuya respuesta solamente requiere del estudiante evocar y repetir o reconocer un contenido trabajado en clase o estudiado. El elemento a recordar puede ser un concepto, una definición, un relato histórico, leyes o principios de una disciplina, un método e, inclusive, la explicación de un fenómeno cuando solo se pide al estudiante que recuerde y reproduzca la explicación estudiada.

Actividades que requieren utilizar o aplicar. Incluimos en esta categoría las preguntas y ejercicios típicamente escolares, cuya resolución requiere estrictamente recordar una fórmula, operación o procedimiento que debe ser aplicado a los datos o información en una situación conocida -y, normalmente, reiterada-, para llegar a una "solución" que suele ser única y predefinida. Se trata de tareas que tienen una única respuesta correcta, a la que se arriba a través de la aplicación de una o varias fórmulas a los datos incluidos en el ejercicio -que suelen, además, ser los estrictamente necesarios, es decir, no faltan ni sobran datos-.

Actividades que requieren construir significados y comprender. Incluimos en esta categoría preguntas y tareas tales como explicar el significado de un concepto; explicar cómo y/o por qué ocurre un fenómeno; explicar un procedimiento, principio o ley. La diferencia con las dos categorías anteriores radica en que se requiere del estudiante construir una explicación que demuestre comprensión. Por ejemplo, una explicación de un hecho histórico que no se limite a reproducir las explicaciones dadas en clase, sino que sea realizada con palabras propias; que el estudiante deba utilizar los conceptos estudiados para explicar un hecho o fenómeno nuevo y distinto de los "dados" en clase. Se incluyen también en esta categoría las tareas que requieren establecer relaciones entre conceptos y/o fenómenos; identificar y extraer información específica de una fuente; observar, registrar y clasificar fenómenos o eventos para construir evidencia; ilustrar o ejemplificar conceptos o principios; realizar inferencias para construir información nueva o extraer

conclusiones a partir de la información dada; fundamentar la respuesta a una pregunta o el procedimiento seguido para resolver un problema.

Actividades que requieren valorar y evaluar. Incluimos en esta categoría a aquellas tareas en las que el estudiante debe evaluar cursos de acción o soluciones alternativas frente a un determinado problema o situación que no tiene una única respuesta correcta; proponer o descartar explicaciones plausibles a fenómenos o problemas que admiten diversidad de respuestas adecuadas; adoptar y argumentar posturas o puntos de vista en situaciones que implican la toma de decisiones basadas en valores; utilizar evidencia empírica con el fin de validar o criticar conclusiones, explicaciones o predicciones; analizar, criticar y/o validar diseños de investigación, conclusiones o procedimientos.

Actividades que requieren diseñar o crear algo nuevo. Incluimos en esta categoría las tareas que implican la utilización del conocimiento disponible para la creación de dispositivos, modelos o procedimientos para resolver situaciones nuevas, complejas y abiertas -problemas en sentido estricto-; la construcción de diseños nuevos en diversos campos -incluidos los diseños para investigar un tema o fenómeno-; la creación de nuevas formas de representar una realidad o fenómeno; y la formulación de predicciones sobre lo que debería ocurrir en una situación dada.

5.1. Recordar y enunciar

En la Figura 2.11 se incluye un ejemplo de actividad de evaluación perteneciente a la primera categoría. Es una pregunta de opción múltiple. El estudiante debe indicar cuál de las cinco posibilidades que se le ofrecen es la respuesta correcta a la pregunta *¿Cuál es el concepto central involucrado en la teoría de Oparin?* Como resulta evidente, para responder solamente se necesita recordar lo trabajado en clase y/o lo leído en un texto. No es necesario ningún tipo de razonamiento, análisis o inferencia.

Figura 2.11: Ejemplo de actividad memorística de múltiple opción en biología

11 - ¿Cuál es el concepto central involucrado en la teoría de Oparin?
a) Formación espontánea de macromoléculas orgánicas
b) Participación de precursores extraplanetarios en el origen de la vida
c) Formación espontánea de moléculas inorgánicas simples
d) Origen de materia no viva a partir de materia orgánica
e) Los eucariotas surgen de los procariotas

Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina**.

La actividad incluida en la Figura 2.12 también pertenece a la categoría "recordar y enunciar". Si bien se presenta como un "ejercicio" e incluye dos ilustraciones, la respuesta a la consigna incluida (que tiene dos partes) no requiere ningún tipo de procesamiento de información, solamente evocar conceptos estudiados. Es bastante evidente que los "sistemas" incluidos corresponden a los estados "sólido" y "líquido", por lo que solo es necesario recordar estas denominaciones. La segunda parte de la consigna puede inducir a creer que se trata de una actividad de nivel superior (correspondiente a la tercera categoría), porque incluye el término "explique". Sin embargo, es claro que no hay nada que explicar. Explicar algo implica que, o bien se está ante un fenómeno nuevo que debe ser explicado, o bien se está ante un interlocutor que desconoce aquello que será explicado. Ninguna de estas dos condiciones se cumple en este caso (se trata de algo ya estudiado que el estudiante debe "explicar" al profesor, que por definición ya conoce el tema). Ante la consigna *"explique las propiedades"* el estudiante deberá reproducir de un modo más o menos fiel lo presentado en el aula por el docente o en las fuentes de estudio empleadas.

Figura 2.12: Ejemplo de actividad de respuesta construida de reproducción en Química

Ejercicio 3: ¿En que estado de agregación se encuentran los siguientes sistemas? Explique las propiedades macroscópicas que presentan.





Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina**.

Un tercer ejemplo de actividad que solo requiere recordar y reproducir es la incluida páginas más arriba, en la Figura 2.5.A. Dicha actividad consta de cinco preguntas relativas a "las hormonas". Se trata también de preguntas "de respuesta construida" (el estudiante no debe seleccionar una respuesta sino producirla por sí mismo). Para responderla el estudiante debe recordar y reproducir conceptos y explicaciones trabajados en clase o estudiadas en un texto[11]. Al responder este tipo de preguntas normalmente los estudiantes se limitan a reproducir, en forma más o menos literal, las explicaciones dadas en clase por el docente o lo leído en

11. Utilizamos las expresiones "texto" para referirnos a una diversidad de fuentes de información: apuntes de clase, fotocopias distribuidas por el docente, libros de texto, internet y toda la amplia gama de fuentes virtuales.

alguna fuente de información escrita. La tarea no incluye información a analizar, ni situación, ni definición de un tipo de producción a alcanzar. Únicamente se pide al estudiante que brinde una respuesta a cada pregunta.

Las tres actividades anteriores tienen en común el hecho de que apelan a un conocimiento *"declarativo"*. Para responderlas el estudiante debe escribir una versión del contenido involucrado que repite o reproduce de manera más o menos fiel lo estudiado.

5.2. Aplicar

Un primer ejemplo de actividad correspondiente a esta categoría es la presentada en la Figura 2.2, que ya fue analizada. En ese caso se trata de una actividad muy simple, que requiere aplicar la división en sucesivas situaciones similares, típicamente escolares. El nivel de requerimiento cognitivo es bajo.

La actividad presentada en Figura 2.4.D es otro ejemplo de esta categoría. Incluye un conjunto de consignas que: i) toman como referencia situaciones cotidianas (desplazarse en un pueblo o ciudad, una médica que administra suero, correr, construir una carretera); ii) cada una plantea dos datos bajo la forma de fracciones (2/3 y 8/12 de kilómetro en el primer caso, litros de suero en el segundo, nuevamente kilómetros en los siguientes); iii) todas terminan con una pregunta cuya respuesta requiere realizar una operación con las fracciones dadas para llegar a un resultado correcto único y predeterminado. La única diferencia que existe entre las tareas es que las dos primeras requieren sumar las fracciones y las dos últimas restarlas. Estas tareas son claros ejemplos de lo que hemos denominado "ejercicios", si bien el docente presenta la actividad bajo la denominación "problemas". Para responder las consignas no es necesario realizar ningún tipo de inferencia ni comprender los fundamentos de las operaciones con fracciones. Probablemente algunos estudiantes sí los comprendan. La mayoría se limitará a aplicar un procedimiento mecanizado. Muchos lo utilizarán en forma equivocada.

La reiteración de tareas con una estructura similar obviamente conduce a mecanizar la utilización del procedimiento, lo cual no es malo en sí mismo, como ya se dijo. La actividad, por otra parte, adolece de una fuerte debilidad en cuanto a su relevancia, por un doble motivo. Primero, porque en las situaciones cotidianas que se plantean no se suelen usar números expresados bajo la forma de fracciones. En segundo lugar, porque la realización "a mano" de operaciones con fracciones es una práctica en desuso en la sociedad, en un tiempo en que las calculadoras están al alcance de todos en los teléfonos celulares. Sobre el análisis de estos aspectos volveremos en el próximo capítulo, al analizar la cuestión de la "autenticidad" de las actividades de evaluación.

En la Figura 2.13 se presenta otro ejemplo de actividad de "aplicación". En este caso la actividad corresponde a un curso de Física en 3º de educación media básica. También en este caso el docente emplea la expresión "problema" para lo que en realidad son "ejercicios".

Figura 2.13: Ejemplo de tareas de respuesta construida de aplicación en física

RESUELVE LOS SIGUIENTES PROBLEMAS

1) ¿Qué tiempo tardará una bicicleta en recorrer 1000 m si su velocidad es de10 m/s?
2) Un atleta demora 14 segundos en recorrer 126 m. ¿Cuál ha sido la velocidad de su carrera?
3) Un móvil recorre medio kilometro en 100 s, ¿qué espacio recorre en...?
4) Un ciclista en su recorrido sufre un aumento de su velocidad de 12 m/s^2 ¿Cuál será la distancia en kilómetros recorrida por el ciclista al cabo de 5 min?
5) Un móvil en su recorrido aumenta su velocidad en $3m/s^2$ ¿Cuál será el espacio en kilómetros recorrido por el móvil al cabo de 8 minutos?

Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina.**

La estructura básica es similar a la analizada anteriormente. La actividad incluye cinco consignas independientes unas de otras. Cada consigna toma como referencia una situación cotidiana (andar en bicicleta, una carrera de atletismo, un "móvil" no definido). Cada consigna incluye dos datos relativos a velocidad, tiempo, distancia o aceleración. Todas las consignas terminan con una pregunta que el estudiante debe responder con un tercer dato que debe "hallar", que es la única respuesta correcta. Para llegar a la misma el estudiante debe utilizar alguna de las fórmulas básicas derivadas de las leyes del movimiento de los cuerpos propuestas por Newton: la velocidad es igual a la distancia recorrida dividida por el tiempo transcurrido; la aceleración es igual a la distancia recorrida dividida por el cuadrado del tiempo transcurrido. Llegar a las respuestas requiere un cierto nivel de inferencia elemental para decidir qué formula utilizar en cada caso, así como el reconocimiento de las unidades de medida básicas y las conversiones entre ellas: metros y kilómetros para la distancia; segundos, horas y minutos para el tiempo; metros por segundo y kilómetros por hora para la velocidad; metros por segundo al cuadrado para la aceleración. Se requiere también un manejo básico de la denominada "regla de tres" para aplicar el principio de proporcionalidad. Finalmente, el estudiante debe realizar adecuadamente un conjunto de operaciones matemáticas.

Resulta evidente que estas tareas tienen cierto grado de complejidad y no son en absoluto sencillas. Sin embargo, no constituyen problemas en el sentido antes definido. Las situaciones planteadas son poco plausibles en la vida real, se trata de situaciones típicamente "escolares". Este hecho, sumado a la reiteración de preguntas con una estructura similar, que se responden con un mismo tipo de procedimiento, llevan a que la mayoría de los estudiantes se acostumbren a que deben "adivinar" qué fórmula utilizar, para luego operar con ella, sin una comprensión cabal de las mismas y de los conceptos físicos involucrados. Algunos estudiantes, los menos, logran comprender los conceptos e interesarse en las actividades. Pero para la mayoría se trata de actividades abstractas, que no llegan a comprender, que intentan resolver por la vía de adivinar qué fórmula se requiere y que, a la larga, terminan generando experiencias de frustración y fracaso.

5.3. Construir significados y comprender

La actividad presentada en la Figura 2.14 es un caso sencillo de tarea que requiere no solamente recordar y enunciar conceptos, sino haberlos comprendido y producir una explicación a partir de una situación cotidiana "auténtica". Para responder las preguntas planteadas el estudiante debe apelar a los conceptos de cambio de estado, calor y temperatura, con el fin de fundamentar, con sus propias palabras, un modo de proceder para resolver una situación de la vida cotidiana. El único requisito para que una actividad como esta sea considerada como perteneciente a esta tercera categoría es que la situación planteada sea efectivamente nueva para los estudiantes. Si la situación ya hubiese sido analizada en clase, entonces la respuesta sería de tipo memorístico y declarativo.

Figura 2.14: Ejemplo de actividad cotidiana auténtica que requiere comprensión

1) Luego de una mañana de intensa lluvia, a la hora del recreo varios alumnos quieren jugar al fútbol. Pero en la cancha hay un gran charco de agua. Andrés, uno de los compañeros propone: *¿Por qué no le pedimos a la señora de servicio un lampazo para extender el agua y que se seque más rápido?*

Analiza la propuesta de Andrés y responde:

a) ¿Estás de acuerdo con la idea de Andrés? ¿Por qué?
b) ¿Cómo puede fundamentarse esa idea desde el punto de vista científico? Utiliza para dicha fundamentación los conceptos sobre cambios de estado, calor y temperatura trabajados en clase.

Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina**.

La actividad presentada antes en la Figura 2.9 es un segundo ejemplo de tarea que requiere de los estudiantes haber comprendido lo estudiado, en este caso, las leyes de la genética. A través de un texto breve la actividad ubica al estudiante en el contexto de una situación auténtica (probablemente real), propia de la investigación en la disciplina. Para responder la pregunta es necesario comprender la situación (en particular las observaciones realizadas por los investigadores en las mariposas), utilizar los datos acerca de las características vinculadas a cada uno de los alelos, recordar las leyes de la genética y relacionar todo esto para inferir cuál de las combinaciones genéticas planteadas da lugar a un determinado tipo de coloración en las mariposas. Para identificar la respuesta correcta en esta pregunta de opción múltiple no alcanza con memorizar las leyes y aplicarlas en forma mecánica. Tampoco es posible descartar las alternativas de respuesta apelando al sentido común.

En la Figura 2.15 se presenta un conjunto de actividades que conforman una prueba de evaluación. Se trata de un caso especialmente interesante, tanto por la situación creada por el docente como por la combinación de tareas de distinto tipo que solicita a sus estudiantes. La situación es difícil de clasificar. Por un lado se trata de una situación ficticia, que toma una serie policial como elemento base. Al mismo tiempo, es posible afirmar que se trata de una situación auténtica, en la medida en reproduce una situación que puede darse en la vida real y en la que los contenidos disciplinares involucrados podrían realmente ser utilizados para responder las preguntas incluidas en las distintas actividades. La situación seguramente es motivadora para los estudiantes y las actividades propuestas incluyen una amplia variedad de información a procesar en distintos formatos (texto, esquema, cifras, ilustración y cuadro).

Las preguntas y tareas propuestas requieren del estudiante una amplia variedad de procesos cognitivos. En algunos casos simplemente se requiere recordar un concepto, como en la pregunta 1.a) "*¿Qué es un sistema?*". La pregunta está seguida inmediatamente de una aplicación del concepto a la situación planteada: "*¿Qué sistema definirías en este caso?*"; "*¿Qué límites utilizarías para delimitar el sistema?*". En otros casos se requiere aplicar fórmulas (ej. 2.a, "*determina el volumen de la habitación*") o aplicar conceptos para la identificación de elementos en una figura (ej. 3.a, "*Indica el alcance, la apreciación y posible estimación de la probeta*"). Simultáneamente, a lo largo de las distintas actividades el docente solicita siempre a los estudiantes que "justifiquen" sus respuestas, es decir, que den razón de las mismas, como forma de demostrar su grado de comprensión de los conceptos y principios que están siendo utilizados.

Figura 2.15: Ejemplo de actividad con formatos diversos que requiere comprensión en Química



Siempre que al Departamento de Policía de Las Vegas se le presentan casos complicados de resolver sin la ayuda de expertos en el área de la criminalística, recurren a científicos que conforman una unidad denominada CSI (Crime Scene Investigation) Ellos se encargan de delimitar sistemas, que denominan escenas del crimen y dentro de ellas analizan las posibles pruebas que se puedan encontrar y que lleven a la resolución del crimen.

El caso que te presentamos a continuación es uno de ellos y tú debes intentar resolverlo. ¡Buena suerte!

El siguiente esquema representa una escena del crimen, donde la víctima fue envenenada con el contenido de una de las tres botellas que están sobre el escritorio:

1.-

a) ¿Qué es un sistema? ¿Qué sistema definirías en este caso?

b) ¿Qué límites utilizarías para delimitar el sistema que definiste? ¿Cómo los clasificarías, por qué? ¿Las paredes de la habitación forman parte del sistema? Justifica

2.- Las dimensiones de la habitación son las siguientes: largo = $4{,}8 \times 10^2$ cm, ancho = $3{,}5 \times 10^2$ cm y altura = $2{,}9 \times 10^2$ cm

a) Determina el volumen de la habitación. Expresa el resultado con el número correcto de cifras significativas.

b) ¿Es lo mismo decir que el volumen de la habitación es 5×10^7 cm^3? Justifica.

3.- Para medir el volumen de líquido que contiene el vaso de la víctima, se utilizó la siguiente probeta:

a) Indica el alcance, la apreciación y posible estimación de la probeta.

b) Determina el volumen que indica la probeta con su correspondiente incertidumbre.

4.-

a) Se sabe que la víctima fue envenenada con el contenido de una de las tres botellas que se encuentra sobre el escritorio. Para saber con cuál de ellas fue, ¿qué tipo de propiedad determinarías? Justifica tu respuesta.

b) Los datos de la investigación se recogen en el siguiente cuadro:

	Botella 1	Botella 2	Botella 3	Vaso
Masa (g) del contenido	900g	600g	900g	150g
Volumen (cm^3) del contenido	800cm^3	600cm^3	750cm^3	125cm^3
Huellas presentes:	Joaquín	José	Julieta	Víctima

Según los datos que puedas extraer del cuadro anterior, determina quién colocó el veneno en el vaso de la víctima. Justifica con cálculos.

Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina**.

Para resolver la actividad final, que consiste en identificar al asesino, los estudiantes deben utilizar buena parte de los conceptos, estimaciones y cálculos realizados en las actividades anteriores. Simultáneamente, deben analizar la información incluida en un cuadro para realizar un conjunto de cálculos e inferencias que también deben ser justificadas.

Esta propuesta de evaluación es un excelente ejemplo de la tercera categoría: comprender y construir significados. Sirve además para mostrar la integración de distinto tipo de procesos cognitivos, que normalmente se

plantean en forma aislada en la mayor parte de las propuestas de evaluación (recordar, definir y aplicar conceptos; realizar cálculos; justificar decisiones y afirmaciones; realizar inferencias), en un conjunto que resulta a la vez novedoso, desafiante y complejo. Se trata propiamente de un "problema", en el sentido definido antes en este capítulo.

5.4. Evaluar

La actividad incluida en la Figura 2.16 puede ser considerada como ejemplo para la tercera, cuarta y quinta categorías de la clasificación según los procesos cognitivos requeridos para su realización.

Figura 2.16: Ejemplo de actividad que requiere comprensión, valoración y creación

El CO_2 atmosférico y el efecto Invernadero

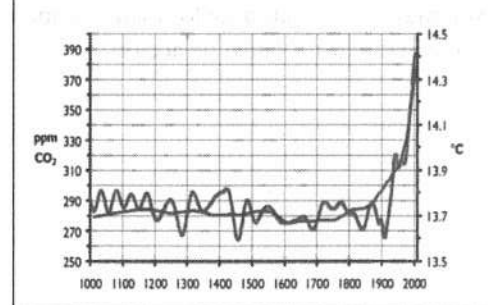
CONSIGNA DE TRABAJO

En clase hemos estudiado qué es el efecto invernadero y cómo éste se relaciona con la presencia de determinados gases en la atmósfera. Uno de estos gases es el dióxido de carbono, CO_2. Algunos científicos afirman que el aumento del dióxido de carbono en la atmósfera producido por las actividades humanas (sobre todo el desarrollo industrial) es el principal responsable del llamado calentamiento global. Otros, sin embargo, dicen que los humanos no somos los únicos responsables, ya qe se trataría de un fenómeno natural.

Suponiendo que tú fueras un periodista que escribe una columna de divulgación científica en un semanario, deberás escribir *un artículo* que se titule: *"Calentamiento global y desarrollo industrial: ¿juntos o separados?"*

Dicho artículo deberá tener las siguientes características:

a. **Contener argumentos a favor y/o en contra de esta relación.**
b. **Incluir la gráfica que se presenta a continuación y hacer referencia expresa a ella en el texto.**
c. **Hacer mención de la Revolución Industrial como hecho histórico en relación con la variación de la concentración de dióxido de carbono en la atmósfera.**



Esta gráfica nos muestra la relación entre la concentración de CO2 (azul) y la temperatura (rojo) durante los últimos mil años

Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina**.

La actividad coloca a los estudiantes ante un fenómeno estudiado en clase, el efecto invernadero, y les recuerda sucintamente algunos de los aspectos trabajados. Se les presenta además información bajo la forma de un gráfico.

La tarea que los estudiantes deben realizar consiste en redactar un artículo de divulgación científica, que debe incluir tres tipos de contenidos, especificados como "características": argumentos a favor y en contra de la relación entre el calentamiento global y los procesos de desarrollo industrial; la información proporcionada por el gráfico; la Revolución Industrial como hecho histórico vinculado con el incremento del dióxido de carbono en la atmósfera.

Para realizar la actividad los estudiantes deben demostrar comprensión de los conceptos involucrados, explicándolos a los lectores del artículo. Es claro, además, que la tarea requiere la construcción de significados a partir de la interpretación del gráfico. En este sentido la tarea corresponde a la tercera categoría. Sin embargo, el contenido principal del artículo es la construcción de argumentos a favor y en contra de una posible relación entre el desarrollo industrial y el calentamiento global. En este sentido es que la actividad puede ser considerada como perteneciente principalmente a la cuarta categoría (que, como se dijo, incluye a las anteriores): implica valorar argumentos a favor y en contra de una determinada afirmación, basándose en información, evidencia y opiniones. Por otra parte, la actividad implica la creación de algo nuevo (un artículo de divulgación), por lo que también podría ser considerada como ejemplo de la quinta categoría.

En la Figura 2.17 se incluye un segundo ejemplo de actividad de evaluación orientada a que los estudiantes valoren cursos de acción a seguir o recomendar, en distintas situaciones vinculadas con los métodos de anticoncepción. Las situaciones planteadas son relativamente sencillas (se definen con un par de frases) y "auténticas". Los estudiantes deben, en cada caso planteado, valorar las ventajas, desventajas y adecuación de los distintos métodos anticonceptivos, con el fin de recomendar un curso de acción.

Figura 2.17: Ejemplo de actividad que requiere valoraciones en Biología

2. Recomienda en cada caso el o los métodos anticonceptivos más eficaces, teniendo en cuenta sus ventajas, desventajas y adecuación

a. Juana tiene 26 años y tiene pareja estable. Lamentablemente por haberse alimentado mal en su adolescencia, sufre de hipertensión.
b. Andrea (29 años) tiene pareja estable, no es fumadora, ni tiene algún problema de salud por ahora.
c. Fabiana (17 años) no tiene novio formalmente pero de vez en cuando se encuentra acompañada de personas que no conoce muy bien.
d. Miguel (38) y Micaela (40) tienen tres hijos y no desean tener más. Micaela ha tenido a lo largo de su vida menstruaciones muy dolorosas.

Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina.**

El elemento que define la cuarta categoría es la evaluación o valoración de argumentos, puntos de vista, situaciones, decisiones o cursos de acción. Es un modo de utilizar el conocimiento que, como se dijo antes, incluye a las categorías anteriores. Es decir, para valorar es necesario recordar y comprender conceptos y principios, así como tener la capacidad para aplicarlos. En muchas situaciones de este tipo los estudiantes deberán, además, recurrir a sus propias convicciones, valores o preferencias y dar cuenta de ellas.

5.5. Diseñar y crear

Tal como se acaba de señalar para la cuarta categoría, también la quinta involucra e incluye a las anteriores. Para diseñar o crear algo nuevo se necesita cierta acumulación previa de conocimientos y habilidades, además de comprender y valorar. La creación es lo que define a esta quinta categoría.

Un ejemplo es la actividad presentada al inicio de este capítulo, en la Figura 2.5.C, que reproducimos en la Figura 2.18. En ella se plantea a los estudiantes una situación propia de la vida social -la construcción de un centro comercial- y requiere de ellos dar forma a un problema y proponer soluciones al mismo. Si bien el planteamiento de la situación es relativamente escueto y simple, claramente se trata de un problema que no tiene una única solución y que implica un conflicto entre fines valiosos: cuidar la flora y fauna autóctonas, construir un espacio de uso público, crear fuentes de trabajo. Ante dicha situación los estudiantes deben proponer cursos de acción desde una perspectiva de desarrollo sustentable.

Figura 2.18: Ejemplo de actividad que requiere creación

2. Analiza el siguiente texto:

"En un área verde, con gran variedad de especímenes de flora y fauna autóctonas, se ha previsto la construcción de un Shopping.

El proyecto implica, por un lado, la destrucción de las tres cuartas partes de la flora del lugar y del hábitat de las especies animales, pero, por otro lado es una fuente importante de mano de obra porque se contrataría gran cantidad de obreros del ámbito local, que, dado el reciente cierre de fábricas han quedado sin trabajo".

a. ¿Cuál es el problema que se plantea en el texto?
b. Desde la perspectiva del desarrollo sustentable, ¿qué posibles soluciones darías al problema? Explica.

Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina**.

El carácter "creativo" de este tipo de actividades depende de dos elementos principales. En primer término, tal como se indicó anteriormente, la situación debe ser realmente nueva para los estudiantes. Al respecto es importante señalar que muchas veces ocurre que la situación ya ha sido discutida o estudiada con anterioridad. En esos casos este tipo de propuestas en realidad se corresponden con la primera categoría (recordar y enunciar), porque los estudiantes se limitarán a repetir lo trabajado en su momento sobre el tema.

En segundo término, el carácter creativo de la actividad depende de los criterios de valoración que el docente utilice para evaluar los trabajos de los estudiantes, que en este caso no están explícitos. Como las consignas incluidas en la actividad son más bien escuetas y no se especifica a los estudiantes cómo deben estar formuladas las soluciones que deben proponer, ni qué grado de detalle y extensión deben tener las respuestas, estas podrían ser muy breves y, en rigor, poco creativas, recurriendo a ciertos "clichés".

La Figura 2.19 contiene un segundo ejemplo de tareas de creación. En este caso los estudiantes deben diseñar tres procedimientos distintos para separar otras tantas mezclas de sustancias. Al igual que en el caso anterior, el carácter creativo de la actividad depende del carácter realmente novedoso de la situación. Es de suponer que en clase han trabajado los conceptos teóricos involucrados y los procedimientos para separar mezclas, pero con otro tipo de sustancias.

Figura 2.19: Ejemplo de actividad que requiere creación

Actividad 3.

Diseña un procedimiento para separar los componentes de las mezclas siguientes e indica qué material de laboratorio necesitarías.

a. Una mezcla de aserrín y arena.
b. Una mezcla heterogénea de tres líquidos inmiscibles: agua, aceite y mercurio.
c. Una mezcla de azufre, limadura de hierro y permanganato de potasio.

Fuente: Bases de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina.**

En el capítulo siguiente, al estudiar con más detalle las características de la evaluación a través de situaciones auténticas, tendremos oportunidad de analizar más ejemplos de actividades de evaluación que involucran la creación y el diseño por parte de los estudiantes.

5.6. Reflexiones finales sobre la clasificación de tareas de evaluación en función de los procesos cognitivos involucrados en su resolución

Estas cinco grandes categorías de actividades de evaluación tienen como finalidad principal servir como herramienta de análisis. Puede ocurrir que para algunas actividades sea difícil decidir a qué categoría corresponden, pero ello no le resta relevancia a la herramienta. La categorización tiene la intención de mostrar la diversidad de tipos de actividades en las que los estudiantes deberían verse involucrados para que el aprendizaje resulte significativo y socialmente relevante.

No pretendemos plantear falsas oposiciones, en el sentido de que cierto tipo de tareas sean buenas y otras malas, o unas necesariamente mejores que otras. En buena medida ello depende de la adecuación a la características de los estudiantes y a los propósitos del curso. De cualquier manera, queremos enfatizar la necesidad de proponer a los estudiantes actividades correspondientes a toda la gama de procesos cognitivos. Nos interesa destacar especialmente la necesidad de no limitarse al uso de actividades de bajo requerimiento cognitivo. No porque sea malo memorizar conceptos, ser capaz de recordar y reproducir, o aplicar procedimientos en forma rápida y mecánica. Por el contrario, son modos de actuar y pensar que tienen valor en sí mismos. Pero no son suficientes, por sí solos, para la formación de personas reflexivas. Tampoco para motivar e involucrar a todos los estudiantes con el trabajo de aprender. Cuando las actividades de enseñanza y de evaluación se limitan a

las dos categorías más básicas y no incluyen los restantes niveles, el aprendizaje pierde sentido para la mayor parte de los estudiantes.

Simultáneamente, es preciso reconocer que los procesos cognitivos más "complejos" requieren de los más "básicos". Existe cierto grado de complejidad creciente y de interrelación entre las capacidades involucradas. La resolución de actividades pertenecientes a las últimas categorías, requieren de procesos cognitivos que incluyen a los correspondientes a las primeras categorías. Dicho en otras palabras, para construir significados es necesario recordar conceptos adquiridos previamente. Para diseñar un dispositivo nuevo es necesario apelar a conceptos o procedimientos más sencillos aprendidos previamente. No es posible evaluar ni crear en el vacío conceptual y procedimental.

Sin embargo, esto no implica que los procesos de pensamiento sean lineales y secuenciales. Desde nuestra perspectiva, es un error pensar que primero se deben memorizar los conceptos y mecanizar los procedimientos, para luego pasar a construir significados, valorar y crear. Consideramos más bien que, en el curso de los procesos de comprender, evaluar y crear, será necesario recordar -o buscar- conceptos y datos, automatizar procedimientos, aplicar reglas. Comprender algo en el marco de una situación significativa y desafiante es, además, una de las mejores formas de memorizar y más tarde recordar algo.

Por tanto, es importante que el lector comprenda que la visión crítica que subyace en este texto hacia las actividades que implican recordar o aplicar, no lo es con relación a las mismas "per se", sino a las prácticas de evaluación y enseñanza que se limitan a ellas y no propician que los estudiantes se enfrenten a actividades de mayor complejidad cognitiva. Dicho en términos más sencillos, no es malo que los estudiantes memoricen conceptos o apliquen procedimientos en forma mecánica, lo malo es que solamente hagan ese tipo de tareas. Lo importante es el panorama que surge al analizar el conjunto de las actividades de evaluación que proponemos a lo largo de un curso. Considerando el conjunto es que cabe hacerse la pregunta de en qué grado estamos propiciando una aproximación reflexiva y crítica a los saberes.

Tampoco debe confundirse "complejidad cognitiva" con una graduación de dificultad. Para muchos estudiantes puede ser más fácil crear algo que recordar datos de memoria cuando los mismos no tienen sentido para ellos, o que resolver ejercicios de física sobre la base de aplicar fórmulas cuyo significado no han comprendido.

Finalmente, es importante advertir que el trabajo con procesos cognitivos complejos debería darse principalmente durante la enseñanza. No es conveniente ni justo con los estudiantes introducir situaciones nuevas, que requieren

valoración y creación, en la instancia de evaluación, si no se ha trabajado en ello durante la enseñanza. Este es el caso de profesores altamente creativos para idear situaciones complejas en las instancias de examen, que no se condicen con el modo en que fue llevado adelante el curso y con el tipo de situaciones que los estudiantes debieron enfrentar y resolver a lo largo del mismo.

6. Resumen y sugerencias para el trabajo en el centro educativo

La enorme mayoría de los docentes queremos una enseñanza más relevante, significativa, reflexiva y motivadora, para los estudiantes y para nosotros mismos. Sin embargo, introducir cambios en nuestros modos de actuar en el aula es un desafío complejo, tal como fue analizado en el primer capítulo.

Un modo concreto de empezar a cambiar nuestras prácticas de enseñanza es comenzar por revisar nuestras propuestas de evaluación. La evaluación es un indicador muy potente del tipo de enseñanza que brindamos. Por otra parte, las propuestas de evaluación en sí mismas son un fuerte determinante de la experiencia educativa de los estudiantes.

> *"Las investigaciones previas sugieren que la naturaleza y forma de las tareas de evaluación ejercen un impacto significativo sobre las disposiciones de los estudiantes hacia el aprendizaje y, muy particularmente, sobre las actividades que llevan a cabo para prepararse para esas situaciones. Los alumnos perciben los requerimientos y exigencias de la evaluación y esas expectativas definen un marco mediante el cual interpretan y reconstruyen el contenido del curso. De este modo, parece ser la evaluación la que conduce el aprendizaje, más que los objetivos establecidos en el curriculum"* (Camilloni y otros, 2010).

Por estos motivos, nuestra intención es proponer un conjunto de herramientas para analizar las actividades de evaluación como camino, primero, para mejorar las propias evaluaciones y, segundo, para revisar en forma más amplia nuestra propuesta de enseñanza. Una primera herramienta, que trabajamos a lo largo de este capítulo, está basada en que toda actividad de evaluación tiene dos componentes centrales: una consigna y un conjunto de información. La consigna o enunciado solicita una respuesta o producción por parte del estudiante. La información le requiere comprender y analizar textos o datos, o lo pone ante una situación. De todas formas, según vimos, la mayor parte de las propuestas de evaluación relevadas en varios países de la región constaban únicamente de una consigna o pregunta, sin información para procesar o situación a resolver.

A partir de dicha distinción inicial analizamos los distintos tipos de consignas que, según su formato, determinan el tipo de tarea o respuesta que el estudiante debe realizar. La principal distinción es entre tareas de respuesta cerrada, construida y abierta. Luego analizamos los distintos tipos de situaciones que se incluyen en las actividades de evaluación, distinguiendo principalmente entre situaciones escolares y situaciones auténticas (así como entre ejercicios y problemas). La clave de esta distinción radica en la similitud de las situaciones que se plantean a los estudiantes con aquellas en las que el conocimiento es producido y utilizado en la vida social y en las disciplinas. Este aspecto será trabajado en profundidad en el capítulo que sigue.

Finalmente, analizamos los distintos tipos de procesos cognitivos que puede requerir la realización de una tarea de evaluación. Vimos que, en un extremo, se ubican aquellas actividades que no incluyen información para analizar (en que la misma es mínima), que proponen situaciones escolares (por lo general reiterativas y artificiales) y que se responden con una frase memorizada, con un dato único o mediante la aplicación mecánica de un procedimiento previamente ejercitado. En el otro extremo se ubican las actividades que implican procesar información abundante y compleja, en situación propias de la vida social y que se responden a través de una producción elaborada, que no tiene una única respuesta ni un modo preestablecido de llegar a ella. Entre ambos extremos existe una amplia gama de posibilidades.

El predominio del aprendizaje entendido como memorización (o aprendizaje superficial) obedece a tradiciones instaladas en la cultura del sistema educativo. Según vimos en el capítulo anterior, estas tradiciones son muy difíciles de modificar, sobre todo si no existen espacios y tiempos institucionales para el intercambio y la revisión sistemáticas en torno a las prácticas de trabajo en el aula.

Sugerencias para el trabajo en el centro educativo

Partiendo de la premisa de que los lectores no se conforman con que sus estudiantes recuerden, repitan y "apliquen" los contenidos que estudian, sino que esperan que sean capaces de comprenderlos, de reflexionar a partir de los mismos y de construir una visión compleja de la realidad, la pregunta central que les invitamos a hacerse, preferentemente trabajando con otros colegas, es: ¿qué tipo de procesos cognitivos promueven las actividades de enseñanza y evaluación que habitualmente proponemos a nuestros estudiantes?

Para transformar esta pregunta en una herramienta de análisis y trabajo colectivo, les proponemos, que cada docente aporte al equipo dos o tres pruebas

o parciales que haya empleado recientemente en sus aulas (3 copias de cada una). Las propuestas se intercambian en el equipo, de modo que cada docente recibe las de dos colegas diferentes, una de cada uno. Las propuestas se distribuyen de modo tal que cada una sea revisada por tres colegas distintos. Ante cada pregunta, ejercicio o tarea concreta incluida en la propuesta, debe responder las siguientes preguntas (teniendo en cuenta lo trabajado a lo largo de este capítulo):

1. ¿Qué formato tiene el tipo de respuesta solicitada? ¿Qué tipo de producción se solicita al estudiante?
2. ¿Es claro y comprensible el enunciado de lo que se espera como respuesta? ¿Qué elementos podrían resultar confusos o ambiguos? ¿Qué cosas se dan por sentadas u obvias y no están explicitadas?
3. ¿Incluye la tarea información a procesar? ¿De qué tipo y qué tan extensa? ¿Es la información comprensible y, a la vez, desafiante?
4. ¿Ante qué tipo de situación se pone al estudiante en esta tarea? ¿Hay una situación? ¿Se trata de una situación escolar o de una situación auténtica? ¿Es disciplinar, personal o social? ¿Es desafiante? ¿Es motivadora? ¿Es "razonablemente" nueva?
5. ¿Qué tipos de procesos cognitivos se requieren para realizar la actividad? ¿En cuál(es) de las cinco grandes categorías definidas se podría incluir la actividad? ¿Por qué? Explicar lo que requiere del estudiante realizar la actividad, teniendo en cuenta lo trabajado a lo largo del capítulo.
6. Formular dos o tres sugerencias para mejorar la propuesta.

Luego del tiempo que sea necesario para responder estas preguntas, se va analizando cada una de las propuestas, escuchando primero la opinión de los "revisores" y luego las reacciones del "autor". Este trabajo de Taller puede requerir varios encuentros.

Referencias bibliográficas

Ausubel, D. y otros (1978). *Educational Psychology: A Cognitive View.* Holt, Rinehart and Winston.

Camilloni y otros (2010). *Los formatos de evaluación de los aprendizajes y sus relaciones con las modalidades de estudio de los alumnos universitarios. Perspectivas de*

investigación y marcos de análisis. Disponible en: http://www.ungs.edu.ar/cienciaydiscurso/wp-content/uploads/2011/11/Camilloni-Basabe-Feeney-20091.pdf. Acceso 24/01/17.

Cohen, L., Manion, L. y Morrison, K. (2004). *Deep and superficial learning.* Publicado en línea como recurso web para "A Guide to Teaching Practice". Disponible en: http://cw.routledge.com/textbooks/0415306752/resources/pdf/10DeepSuperficialLearning.pdf. Acceso: 20/05/16.

Leymonié, J. y otros (2013). *La evaluación en las aulas de Ciencias Naturales en cuatro países de América Latina: Concepciones de Ciencia y de Educación Científica.* Documento no publicado que forma parte de los productos del estudio "La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina".

Lublin, J. (2003). *Deep, Surface, and Strategic Approaches to Learning.* Documento publicado en línea por el Centre For Teaching and Learning de Dublin. Disponible en: http://www.ed.ac.uk/institute-academic-development/learning-teaching/staff/advice/researching/publications/experience-of-learning. Acceso: 20/05/16.

Marton, F. y Säljö, R. (1976). *On qualitative differences in learning. 1 - Outcome and process.* En: British Journal of Educational Psychology, Nº 46, 4-11.

Marton, F. y Säljö, R. (2005). *Approaches to learning.* En: Marton, F., Hounsell, D. and Entwistle, N., (eds.), The Experience of Learning: Implications for teaching and studying in higher education. 3rd (Internet) edition; pp. 39-58. Edinburgo: University of Edinburgh, Centre for Teaching, Learning and Assessment. Disponible en: http://www.ed.ac.uk/institute-academic-development/learning-teaching/research/experience-of-learning. Acceso 24/01/17.

Piaget, J. y otros (1973). *La representación del mundo en el niño.* Madrid: Ediciones Morata. (Original en inglés publicado en 1926).

Picaroni, B. y Loureiro, G. (2010). *Qué matemática se enseña en las aulas de sexto año de Primaria en escuelas de Latinoamérica.* En: Páginas de Educación, Vol. 3, Año 3. Montevideo: Universidad Católica del Uruguay.

Ravela, P. y otros (2014). *La evaluación en las aulas de secundaria básica en cuatro países de América Latina.* En: Propuesta Educativa, Año 23 Nº 41, pp. 20 a 45. Buenos Aires: FLACSO.

Säljö, R. (1979). *Learning in the learner's perspective. I. Some common-sense conceptions.* Reports from the Department of Education, University of Goteborg, 76.

Stone Wiske, M. (comp.) (1999). *La enseñanza para la comprensión. Vinculación entre la investigación y la práctica.* Buenos Aires: Paidós.

Van Rossum, E. J. and Schenk, S. M. (1984). *The relationship between learning conception, study strategy and learning outcome.* British Journal of Educational Psychology, 54, pp. 73-83.

Wiggins, G. (1998). *Educative Assessment. Designing Assessments to Inform and Improve Student Performance.* San Francisco: Jossey-Bass.

CAPÍTULO 3

EVALUAR A TRAVÉS DE SITUACIONES AUTÉNTICAS

Introducción

En el capítulo anterior nos enfocamos en el análisis de las actividades de evaluación desde la perspectiva de los procesos cognitivos que las mismas requieren de los estudiantes. Desarrollamos una primera herramienta para la revisión de las actividades que utilizamos habitualmente en el aula para evaluar aprendizajes, analizando los tipos de tareas que requerimos realizar a los estudiantes, la información que deben analizar y los procesos cognitivos que deben poner en juego para ello. Utilizamos como marco conceptual la distinción entre aprendizaje superficial y aprendizaje en profundidad, establecida a partir de los años 70 por investigadores suecos. Mostramos que esta distinción está emparentada con otros marcos conceptuales, como el aprendizaje significativo y la enseñanza para la comprensión.

En términos generales, mostramos que la mayoría de las actividades de evaluación registradas en el marco de nuestros trabajos de investigación en varios países de América Latina[1] son de bajo requerimiento cognitivo. Predominan las actividades breves, que casi no requieren del estudiante analizar información, sino más bien recordarla y reproducirla. Predominan también las actividades en que el estudiante debe utilizar una fórmula o un algoritmo matemático para resolver ejercicios que tienen una única respuesta correcta y que suelen presentarse como series de ejercicios similares. Son escasas las actividades que requieren del estudiante demostrar una verdadera comprensión del contenido con el que trabaja y, menos frecuentes aún, las que implican valorar situaciones, diseñar o crear nuevos dispositivos, o resolver situaciones nuevas propias de la vida real o de producción de conocimiento en la disciplina.

1. Las referencias bibliográficas y la descripción de estos trabajos fueron incluidas en el capítulo 1.

Como parte de la herramienta de análisis de las actividades de evaluación, les propusimos observar los tipos de situaciones que las mismas plantean a los estudiantes. Mostramos que hay situaciones que denominamos "auténticas" (personales, sociales o disciplinares), en las que el conocimiento se produce y utiliza de un modo similar al que tiene lugar en la vida real. En tanto que denominamos como "escolares" a todas aquellas situaciones que difícilmente el estudiante encuentre fuera del ámbito escolar. Estas últimas no reflejan adecuadamente las formas de producir y utilizar el conocimiento en la vida real, solamente tienen lugar en el ámbito escolar y, por tanto, le dan un carácter de artificialidad y falta de sentido a los contenidos de la enseñanza. Este tipo de situaciones hacen más difícil que el estudiante tenga encuentros significativos con dichos contenidos. Como consecuencia, el aprendizaje se torna poco motivador y rutinario.

Este capítulo tiene como propósito profundizar en la comprensión de qué son situaciones auténticas, por qué son importantes y cómo crearlas. A lo largo del mismo les invitaremos a dar un paso más en la revisión de las actividades de evaluación, desde la perspectiva de la situación en la que ponen a los estudiantes. La pregunta principal que nos haremos es cómo construir situaciones que sean intelectualmente desafiantes y que no sean artificiales, sino similares a aquellas en que el conocimiento es creado y utilizado en las disciplinas y en la vida ciudadana. Para ello analizaremos el enfoque de la "evaluación auténtica", surgido a finales de los años ochenta y principios de los noventa como parte de un conjunto más amplio de iniciativas pedagógicas orientadas a dotar de mayor significación y sentido al aprendizaje.

1. La evaluación auténtica

El movimiento de la evaluación auténtica surge como modelo alternativo a las formas de evaluar predominantes en los sistemas educativos a lo largo del siglo XX, las que estuvieron orientadas principalmente a la memorización y reproducción declarativa de los contenidos, así como a la "resolución" rutinaria de ejercicios repetitivos. Como norma general estas evaluaciones constituyen la etapa final del proceso de enseñanza. Suelen ser planteadas al finalizar un tema, ciclo o módulo y estar conformadas por tareas y preguntas descontextualizadas, que solo tienen sentido en el ámbito escolar. Rara vez se permite a los alumnos planificar sus producciones antes de la instancia de evaluación, ni elaborar una versión inicial para luego revisarla y rehacerla a partir de una devolución o revisión.

La expresión "evaluación auténtica" fue acuñada por Grant Wiggins a fines de los 80, según él mismo lo indica en un "post" del 24 de enero

de 2014 titulado *"Authenticity in assessment, (re-)defined and explained"* (*La autenticidad en la evaluación redefinida y explicada*)[2]. En esta publicación Wiggins recuerda que el término fue acuñado mientras trabajaban en un movimiento de reforma educacional denominado "Coalición de Escuelas Esenciales", liderado por Theodore Sizer. La intención original, afirma, fue distinguir entre pruebas "verdaderas", por oposición a las meramente académicas, irreales y artificiales que predominaban en ese momento (y lo siguen haciendo[3]).

La definición original de evaluación auténtica de Wiggins fue la siguiente[4]:

> *"Pruebas auténticas son desafíos representativos de las tareas propias dentro de una disciplina determinada. Son diseñadas para enfatizar un grado de complejidad realista (pero, a la vez, justo y razonable); enfatizan la profundidad más que la amplitud. Para hacer esto, necesariamente deben involucrar tareas o problemas poco estructurados y que tengan cierto grado de ambigüedad"* (Wiggins, 1989).

Según explica el propio Wiggins, la primera utilización de la expresión "evaluación auténtica" en un libro, fue autoría de su colega Fred Newmann. El título del libro era "*Más allá de la evaluación estandarizada: evaluando el logro académico auténtico en las escuelas secundarias*" (Archbald y Newmann, 1988). La definición propuesta por Newmann en ese trabajo y otros posteriores, es la siguiente: *"El logro intelectual auténtico está basado en tres criterios: la construcción de conocimiento por parte del estudiante* [1], *a través de la indagación disciplinar* [2] *que tiene valor más allá del salón de clases* [3]*"*[5].

---

2. https://grantwiggins.wordpress.com/2014/01/26/authenticity-in-assessment-re-defined-and-explained/. Acceso el 27/05/16 (justo en el primer aniversario de su muerte repentina).
3. Especialmente en el contexto de los Estados Unidos, en el que había y sigue habiendo una enorme cantidad de pruebas estandarizadas de respuesta cerrada, utilizadas para todos los fines (diagnóstico, calificación de estudiantes, selección para la universidad, entre otros).
4. En el original, en inglés: *"Authentic tests are representative challenges within a given discipline. They are designed to emphasize realistic (but fair) complexity; they stress depth more than breadth. In doing so, they must necessarily involve somewhat ambiguous, ill structured tasks or problems"*.
5. En el original, en inglés: *"Authentic intellectual achievement based on three criteria: student construction of knowledge through disciplined inquiry that has value beyond the classroom"*. Avery, Patricia. Authentic Assessment and Instruction. Publicado en línea en: http://www.socialstudies.org/system/files/publications/se/6306/630612.html. Acceso 27/05/16.

El foco de la preocupación de estos autores es el carácter artificial y poco realista de las formas de relación con el conocimiento que se promueve en las instituciones educativas. La tendencia predominante es presentar a los estudiantes formas "escolarizadas" del conocimiento. Este no aparece como una construcción humana compleja, inacabada, en evolución continua y vinculada con proyectos y necesidades diversos de la vida personal y social, sino como algo preestablecido, fijo y desvinculado del contexto histórico y social, que debe ser aprehendido, memorizado y repetido en las instancias de evaluación, en términos más o menos similares a aquellos en los que fue presentado en clase o en los libros de estudio.

Los impulsores de la evaluación auténtica proponen evaluar -y, por tanto, también enseñar- a través de situaciones que sean más parecidas a aquellas en las que el conocimiento es producido y utilizado en la vida real, tanto en la sociedad en general como por parte de quienes se dedican a la construcción del mismo en las diferentes disciplinas. De allí la preocupación de Wiggins, en la definición antes citada, por pruebas que constituyan *"desafíos representativos de las tareas propias de una disciplina"*, que tengan cierto *"grado de complejidad"* y que constituyan *"problemas poco estructurados"*. En el caso de Newmann, el énfasis está puesto en la actividad del estudiante construyendo el conocimiento e indagando en torno a tareas que tengan sentido más allá del aula. Esto último implica romper con las prácticas de evaluación fuertemente instaladas, que le requieren al estudiante escribir para el docente sobre lo que el docente ya sabe.

Según es posible apreciar, el enfoque está claramente vinculado con la distinción entre "ejercicios" y "problemas" propuesta en el capítulo anterior, así como con la distinción entre "aprendizaje superficial" y "aprendizaje en profundidad".

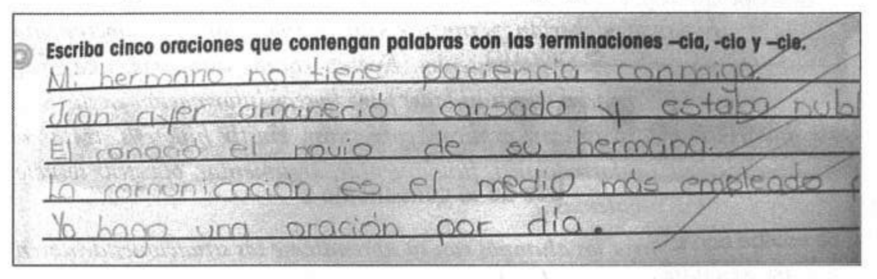
1.1. Iniciativas orientadas a la enseñanza y a la evaluación auténticas en la educación latinoamericana

Esta preocupación por la autenticidad en la evaluación y en la enseñanza no fue inspiración exclusiva de estos autores, sino que se produce en forma más o menos paralela -a veces conectada- con otras iniciativas educativas que surgieron en distintas partes del mundo en la segunda mitad del siglo pasado y que van en la misma dirección. Mencionaremos, a título ilustrativo y con el fin de ampliar la visión del lector, algunas de dichas iniciativas que emergieron en América Latina.

Una de estas nuevas aproximaciones en búsqueda de una enseñanza más "auténtica" está constituida por los trabajos pioneros de Emilia Ferreiro

(1982, 1986, 1991, 1998) con relación a la enseñanza del lenguaje. Ferreiro dio un fuerte impulso a la enseñanza de la lengua escrita a través de actividades con sentido social: escribiendo para un destinatario real y sobre cosas que tuvieran significado para los alumnos. Se intentó desterrar de la escuela las antiguas "redacciones" del tipo "Mis vacaciones" y el aprendizaje de la escritura a través de frases mecánicas con poco sentido ("Mi mamá me ama") o de palabras con una misma letra (la-le-li-lo-lu). No obstante, como lo muestra la actividad incluida en la Figura 3.1, este tipo de prácticas persiste aún en 6º grado de Primaria, en algunos países de nuestra región.

Figura 3.1: Actividad de evaluación en lenguaje en 6º de Primaria


Escriba cinco oraciones que contengan palabras con las terminaciones –cia, -cio y –cie.
Mi hermano no tiene paciencia conmigo.
Juan ayer amaneció cansado y estaba nubl
El conoció el novio de su hermana.
La comunicación es el medio más empleado
Yo hago una oración por día.

Fuente: Base de registros del estudio **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas.**

El enfoque de Emilia Ferreiro puso el énfasis en la función comunicacional del lenguaje, por lo cual el aprendizaje del mismo debe realizarse a través de actividades que tengan realmente esta función: escribir cartas con un destinatario real; escribir en un periódico escolar; participar en concursos de cuentos que luego se "publican" para las familias; preparar, escribir y realizar representaciones teatrales, entre otras. Se buscó de este modo reemplazar el aprendizaje mecánico de la escritura y el foco excesivo en contenidos gramaticales abstractos, sin relación con la escritura "real" de los estudiantes. Dicho en otras palabras, la propuesta de Ferreiro consistió en enseñar el lenguaje a través de actividades auténticas, en lugar de hacerlo a través de situaciones escolares carentes de sentido.

Este enfoque se desarrolló también en otras disciplinas, especialmente en la matemática y las ciencias naturales. En estos ámbitos se pone hoy especial énfasis en "hacer" matemática y "hacer" ciencias. No se trata de aprender de memoria los pasos del método científico, ni contenidos científicos en forma abstracta y descontextualizada, sino de aprender a pensar científicamente, a ser usuarios reflexivos del conocimiento científico en situaciones de la vida social y, eventualmente, a "construir ciencia" como lo hacen quienes se

dedican a ella (con las diferencias obvias e ineludibles, por razones de edad, acumulación de conocimiento y condiciones materiales).

En este sentido Beatriz Macedo, especialista de la UNESCO en Educación Científica, expresaba en el año 2006 en una ponencia presentada en el Congreso Internacional Didáctica de las Ciencias realizado en Cuba:

> *"Integrar la cultura científica como parte indisociable de la cultura supone asegurar desde las edades más tempranas una educación científica adecuada. Ello nos hace insistir en la necesidad de proponer tareas que involucren a los niños en su aprendizaje, a partir de la presentación de situaciones problemáticas, con temas de su interés, que sean próximos a sus realidades y que posean relevancia social para su entorno y su país.*
>
> *Estas situaciones deberían permitirles interactuar con sus conocimientos previos y con sus ideas preexistentes. Es importante, asimismo, que ofrezcan muchas y variadas oportunidades para buscar información en distintas fuentes, reunir, clasificar, explicar, conjeturar, emitir hipótesis, tratar de buscar nuevas informaciones, fundamentar, argumentar, plantear soluciones y nuevos problemas...*
>
> *Para involucrar a los alumnos con su aprendizaje las situaciones de aprendizaje presentadas en el aula deben estar íntimamente conectadas con las necesidades sociales, pertenecer a la realidad inmediata del alumno y relacionarse con los avances técnicos de los cuales la mayoría de los ciudadanos somos usuarios. Bajo esta óptica la enseñanza de las ciencias se convierte en un instrumento para la alfabetización científico-tecnológica de los ciudadanos, que los ayuda a comprender los problemas que tiene la sociedad actual y los faculta para la toma de decisiones fundamentadas y responsables"* (Macedo, B. 2006).

En la misma línea, haciendo referencia al concepto de alfabetización científica de las pruebas PISA (OCDE, 2006), Leymonié y otros (2013: 4) señalan:

> *"En los últimos 50 años los educadores en ciencias han estado interesados en los nuevos escenarios socio-culturales donde se destacan los problemas ambientales y la preocupación por la calidad de vida. Evidencia de tal interés han sido los cuestionamientos a los objetivos de la formación científica: se reclaman nuevas alfabetizaciones respecto de los contenidos de ciencias y los modelos de prácticas educativas... Existe hoy un consenso claro respecto a la necesidad de que la alfabetización científica se imponga como una dimensión esencial de la cultura ciudadana. Los educadores científicos han arribado a la idea de que una educación que tenga como propósito la alfabetización científica debería incluir como contenidos los conocimientos*

y procedimientos de la Ciencia (datos, hechos, conceptos, teorías, técnicas, etc.) así como también los conocimientos sobre la Ciencia (historia y naturaleza de la Ciencia, modelos, explicación, método, etc.) Todo esto en un marco de aplicación a la resolución de problemas situados en contextos reales que exija pensar también sobre los aspectos éticos, sociales y económicos de los temas científicos y tecnológicos".

Estos aspectos han estado presentes, además, en el debate curricular de las últimas décadas en varios países de la región, especialmente en torno a los contenidos de la educación media, tradicionalmente orientados a la preparación de los estudiantes para la Universidad. Uruguay es un caso paradigmático de esta situación. La educación secundaria en este país fue creada inicialmente como "sección preparatoria" de la Universidad de la República, de la cual formaba parte. En el año 1935 se independizó de la misma, pero mantuvo su impronta curricular orientada a saberes propios de la actividad universitaria, trabajados además de un modo abstracto y declarativo[6].

Esta situación dio lugar a varios intentos fracasados de reforma de dicho nivel educativo a lo largo de los últimos cincuenta años, el primero de ellos conocido como *"el Plan 63"* (por el año 1963), en el que se produjo una fuerte reacción contra el enciclopedismo del Plan de Estudios vigente, elaborado en el año 1941, orientado a la formación específica del estudiante preuniversitario, con énfasis en actividades académicas valoradas a través de calificaciones. La reforma del 63 surgió con un espíritu de formación integral para la vida, el trabajo, la salud, el hogar y la recreación, intentando vincular a la instituciones de educación media con las comunidades en las que estaban insertos los estudiantes, mediante el planteo de problemas concretos que eran resueltos en talleres. Incluía además un innovador sistema de evaluación de los desempeños de los estudiantes (ANEP, 2002).

En una entrevista realizada en el año 2014, el Profesor Glauco Daniel Cabrera, participante directo de la experiencia, brinda los siguientes testimonios:

"El Plan entonces vigente (1941)... funcionaba más bien como... antesala de los estudios universitarios. Estudios, métodos y exámenes fuertemente marcados por una nítida orientación profesionalizante: asignaturas cerradas sobre sí mismas y propuestas como «monumentos escritos» a amontonar

6. Esto, que pudo haber tenido sentido en la primera mitad del siglo XX, cuando solo una élite accedía a la enseñanza secundaria, ha dejado de tenerlo en el siglo XXI, cuando la educación formal ha sido legalmente declarada como un derecho y una obligación hasta el final de la Educación Media Superior.

unos sobre otros y a memorizar y a reproducir tal como figuraban en numerosos manuales...

Postulado el principio de actividad en sus múltiples dimensiones [se refiere al Plan 63], *como vector principal del desarrollo de la personalidad, las temáticas diversas se proponen como no taxativas, facilitando así la flexibilidad y la adaptación a las diferentes situaciones y relaciones pedagógicas reales, en cada sala de clase, cada día. Se brindaba de esta manera a profesores y alumnos, la posibilidad concreta de construir proyectos de trabajo, combinando exigencias de aprendizaje con intereses inmediatos y mediatos, con problemáticas renovadas, con preguntas abiertas. De hecho, los enfoques lineales y cerrados de los campos disciplinarios se desdibujan y dan lugar a la puesta en práctica de "unidades de trabajo" pluridisciplinarias que hacen de cada jornada en el liceo una sucesión de tareas de investigación y de descubrimientos...*

Otra particularidad destacable de la organización curricular la constituye la participación de los alumnos en Talleres de Manualidades y en Actividades Facultativas. Durante el Primer Nivel del Primer Ciclo, todos viven la experiencia del trabajo manual en todas sus etapas, desde la concepción hasta el resultado final. Orientados siempre por profesionales de reconocido nivel y con gusto y capacidades de transmisión. Subrayemos la igualdad estatutaria de estas "prácticas" respecto a todas las otras actividades: tan formativo es aprender a desarmar un motor que comprender un poema de Góngora o intentar un "collage" que interpretar una pintura de Picasso. En el mismo espíritu pueden elegir, en una lista confeccionada colectivamente, actividades que incluyen una muy amplia gama de posibilidades, en diversos órdenes desde organizar eventos de diversión o entretenimiento hasta aprender a discutir y argumentar en una asamblea bien organizada, etc." (García Castro, A. 2014).

En este marco curricular y pedagógico, la evaluación también era realizada con un enfoque innovador para la época:

"La evaluación se realiza mediante la redacción obligatoria de una «monografía» sobre un tema o problema elegido por cada alumno a partir de sus cursos y orientado por sus profesores. Acompañado individualmente en sus investigaciones (bibliotecas, archivos, terreno) por un "profesor-guía" debe presentar planes y textos en plazos preestablecidos y controlados. El enfoque, obligatoriamente pluridisciplinario. La calidad del trabajo sería juzgada por el equipo de docentes. Si resultaba satisfactoria el alumno "obtenía" su NCB. Si el texto presentaba limitaciones, deficiencias o lagunas, debía ser

redactado nuevamente y su nueva versión "defendida" delante de un jurado especial" (García Castro, A. 2014).

El desarrollo de esta experiencia se vio interrumpido por el período autoritario que vivió el país entre los años 1973 y 1984. Intentos posteriores de reforma curricular en la educación secundaria, realizados en 1995 -en la secundaria básica- y 2003 -en la secundaria superior- terminaron también fracasando. No es casual que hoy Uruguay ostente uno de los índices más bajos de finalización de la educación media superior en la región: apenas el 40% de los jóvenes la terminan (INEEd, 2014), a pesar del temprano desarrollo de su sistema de educación primaria, que alcanzó una cobertura total en los años 60.

Las tres propuestas educativas brevemente reseñadas (de enseñanza del lenguaje escrito, de educación científica y de reforma curricular en la educación media) son apenas algunas muestras de la preocupación y las iniciativas en la región, en torno a la necesidad de una educación orientada hacia formas de relacionamiento con los saberes que sean personal y socialmente más relevantes. Por cierto, existen muchas otras que el lector seguramente conozca.

1.2. Enseñanza para la comprensión vs. enciclopedismo: la polémica Gardner-Hirsch

Volviendo a los autores originales de la expresión "evaluación auténtica" (Wiggins y Newmann), es importante señalar que ambos se inspiran en los trabajos y aportes de Gardner a través de su propuesta de enseñanza para la comprensión. El eje central de la misma, según explicamos en el capítulo anterior, es el propósito de lograr que los estudiantes comprendan los contenidos y no solo los memoricen. El movimiento de la enseñanza para la comprensión surgió en los Estados Unidos como contrapropuesta al excesivo énfasis en la evaluación de los estudiantes a través de pruebas estandarizadas con actividades simples y de tipo memorístico.

En la década de los 90 se produjo un muy interesante debate entre Howard Gardner y Eric Hirsch en torno a este tema. Hirsch, defensor de la evaluación estandarizada, había publicado a fines de los 80 un libro titulado "*Alfabetización Cultural. Lo que todo americano necesita saber*"[7]. Su argumento principal es que los niños y jóvenes necesitan "saber" de muchos temas

7. El título original en inglés es *Cultural Literacy. What Every American Needs To Know*. Fue publicado en 1989.

y tener un "contexto cultural" para poder pensar. Por ejemplo, deben saber cuándo fue la guerra civil en los Estados Unidos y qué es un chip de silicona. Desde esta perspectiva, Hirsch propone una lista de cinco mil datos, nombres, frases, fechas, lugares y conceptos que todo estudiante debería aprender en la escuela para llegar a ser un ciudadano pleno. Defensor de las pruebas estandarizadas y de los métodos educativos tradicionales, publicó a lo largo de los años 90 una serie de libros dedicados a establecer lo que los niños deberían aprender en cada grado escolar (un libro por cada grado)[8].

A esta propuesta, de corte claramente enciclopedista (en el buen sentido del término), Gardner le respondió con otro libro, titulado "*La mente disciplinada. Lo que todo estudiante debería comprender*"[9]. La expresión "disciplinada" no refiere a disciplina en el sentido de comportamiento u orden, sino en el sentido de campo del saber: ciencias, artes, etc. El propio título del libro, "lo que todo estudiante debería comprender", expresa la intención del autor de confrontar entre "saber" y "comprender". Con esta finalidad, el libro está orientado a defender la importancia de enseñar menos temas con mayor profundidad. Como forma de provocar al lector y de ilustrar su perspectiva, Gardner explica a lo largo del libro cómo se podría confeccionar un currículo para un año entero de enseñanza, en torno a tres temas cuidadosamente elegidos: la Teoría de la Evolución, el Holocausto del pueblo judío y la ópera Las Bodas de Fígaro. A lo largo del libro Gardner muestra la complejidad de cada uno de estos temas y todas las posibles actividades educativas que podrían realizarse en torno a ellas. Y, en particular, desarrolla la importancia de aprender a pensar en profundidad y en forma "disciplinada". Para ello es necesario elegir pocos temas ("menos es más") y trabajarlos con la mayor profundidad posible, ya que la educación enriquece a los individuos en la medida en que los ayuda a comprender el mundo (Gardner, 1999).

Lo importante de la propuesta de Gardner no son estos tres temas en sí mismos, sino su carácter representativo de las tres grandes formas de pensamiento disciplinar que deberían desarrollar los estudiantes: el vinculado a la verdad y a las construcciones científicas (la Teoría de la Evolución); el vinculado a la ética, la moral y la historia humana (el Holocausto del pueblo judío) y el campo de lo estético y la creación artística (las Bodas de Fígaro).

En forma irónica Gardner afirma que "*la alfabetización cultural, con su promesa de cinco minutos para cada tema, parece más atractiva que el conoci-*

8. Por ejemplo, *What your first grader needs to know* (que traducido sería *"Lo que su hijo de primer grado necesita saber"*) publicado en 1991.
9. El título original en inglés es *The Disciplined Mind. What all students should understand*, publicado en 1999.

miento en profundidad de una inevitable selección de temas. Sin embargo, en ausencia de un modo de pensar disciplinar, la alfabetización cultural carece de un "hogar epistemológico"; amontona una mezcolanza de conceptos y datos ('bien alumnos, ya fue suficiente sobre el Holocausto, pasemos ahora a los Hologramas') que esperan ser usados de alguna manera, en algún momento, en algún lugar. Más aún, en ausencia de una textura y 'pegamento' disciplinar, los datos se olvidan fácilmente" (Gardner, 1999: 118).

El objetivo principal de la educación no es conocer superficialmente muchas cosas, sino aprender a pensar y trabajar en forma disciplinada en las principales áreas de la creación cultural humana. "*Es más probable que las personas aprendan a pensar como científicos si se involucran profundamente en un área (como las causas del cáncer, de la pobreza o del estrés) que si sobrevuelan sobre cien ejemplos distintos tomados de una docena de ciencias*" (125).

En el contexto de este enfoque, Gardner desarrolla una propuesta de evaluación de aprendizajes diferente a las prevalecientes en su país en ese momento -y que aún persisten-. La idea central es lo que denominó "desempeños de comprensión". La evaluación debería realizarse principalmente a través de exhibiciones públicas de lo que los estudiantes han aprendido y son capaces de hacer. Por ejemplo, predecir qué pasaría a una determinada especie si se produce un cambio drástico en la ecología local; la creación de una canción lírica que capture las diferencias generacionales en la sociedad contemporánea; o un análisis de una lucha violenta entre grupos étnicos, en términos de sus similitudes y diferencias con el Holocausto (131). En vez de estar sometidos a exámenes secretos, los estudiantes deberían conocer de antemano qué tipo de desempeño se les va a pedir y con qué criterios serán evaluados. Deberían poder incluso discutir dichos criterios. Y deberían tener tiempo para preparar su presentación (denominada por el autor "exhibición" o "performance").

Por otra parte, la evaluación debería ser continua en lugar de ocurrir al final de un curso o unidad. A medida que elaboran sus desempeños, trabajos o exhibiciones, los estudiantes deberían recibir devoluciones y sugerencias de sus docentes y de sus compañeros. Todos los estudiantes deberían saber que tendrán oportunidades para mejorar, completar y lograr su "desempeño de comprensión", dado que así es como se trabaja normalmente en la vida real y en la producción de conocimiento.

1.3. Las situaciones realistas como clave de la evaluación auténtica

Los trabajos de Gardner y de los grupos que han trabajado en el enfoque conocido como "enseñanza para la comprensión" están en la base de

los desarrollos de la "evaluación auténtica". Sin embargo, como muestran Frey y Schmitt (2012) en un artículo en el que revisan la diversidad de conceptos y definiciones que sobre el tema se desarrollaron a lo largo de veinte años, el concepto de evaluación auténtica fue ampliando y diversificando su significado. Se la asocia con tres elementos diferentes: i) el contexto de la evaluación, que debe ser realista, requerir la demostración de un desempeño y constituir una tarea cognitivamente compleja; ii) el papel del estudiante, que debe suponer la defensa de un producto e implicar formas de trabajo colaborativo y evaluación formativa; y, iii) el modo de calificar al estudiante, que debe incluir criterios claros y conocidos, el uso de múltiples indicadores y, en algunos casos, el uso de instrumentos específicos como los portafolios.

Dado que algunos de estos elementos no son exclusivos de la evaluación auténtica, los autores proponen enfocar la definición en lo central, propio y específico de la misma, según fue desarrollado por Grant Wiggins: una actividad de evaluación que involucra profundamente al estudiante, que está enfocada en capacidades y situaciones que tienen valor más allá de la propia evaluación y del aula, y que tiene carácter "realista". Según Wiggins, la evaluación es auténtica cuando las actividades que proponen son similares a las actividades que llevan adelante las personas en la vida real (1998: 21).

Para comprender mejor en qué consiste la propuesta conviene exponer algunos ejemplos. El primero de ellos está incluido en el libro de Wiggins antes referido (1998: 27). Se trata de una actividad de matemática, dirigida a estudiantes de 8º grado (Recuadro 3.1).

Recuadro 3.1: La sección de envolver regalos en una tienda

La situación:

Estás a cargo de la sección de envolver regalos de una tienda. En promedio, 24.000 clientes compran ropa en la tienda cada año. Alrededor del 15% de los clientes piden que su compra sea envuelta para regalo. Cada mes se venden aproximadamente 165 chaquetas, 750 camisas, 480 pantalones y 160 sombreros. El precio de todas las cajas para ropa es el mismo. El metro de papel de regalo cuesta cinco pesos. Cada rollo de papel de regalo tiene un metro de ancho y 100 metros de longitud.

Como responsable de la sección necesitas planificar los costos para el año y, obviamente, quieres ahorrar dinero lo más posible. ¿Cuál sería la forma de las

cajas para pantalones, camisas, chaquetas y sombreros que requeriría la menor cantidad de papel de regalo?

Tu tarea: Recomendar a la sección de compras, en un informe escrito:

- El tamaño y la forma de las cajas que deberían comprarse para los pantalones, camisas, chaquetas y sombreros, cuando estos son comprados por separado.
- La cantidad de rollos de papel de regalo que se necesitan para un año.
- El costo aproximado del papel de regalo que se necesita para las ventas de un año de pantalones, chaquetas, camisas y sombreros.

Cuestiones a tener en cuenta:

1. ¿Qué tamaño tienen que tener las cajas para poner en ellas la ropa doblada? El modo en que se doble la ropa hace una diferencia en el tamaño de la caja que se requiere.
2. Experimenta midiendo, doblando y poniendo ropa en las típicas cajas de cartón liviano que se utilizan en las tiendas (o construye cajas con cartulina para experimentar).
3. Ten en cuenta que ciertas formas de caja son más fáciles de envolver que otras, con mínimo desperdicio. Pero las formas con las que se desperdicia menos, pueden requerir más cantidad de papel en el envoltorio. ¿Puedes pensar en alguna regla o generalización acerca de la cantidad de papel de regalo que requiere una determinada forma de caja en relación con la cantidad de desperdicio de papel que podría evitarse usando una caja más grande?
4. No es posible envolver un paquete sin desperdiciar algo de papel. Piensa en el costo del papel extra y en el papel que se deja de usar de cada rollo, dados los requerimientos de la tarea en el mundo real.

Tu trabajo será evaluado en relación a los siguientes criterios:

- La sofisticación matemática de tu propuesta.
- Los razonamientos y métodos matemáticos.
- La efectividad de la propuesta.
- La calidad del informe.
- La precisión de la propuesta.

Fuente: Wiggins, G. (1998: 27).

La tarea propuesta constituye un ejemplo prototípico de actividad auténtica. Se plantea a los estudiantes una situación concreta, enmarcada en un contexto propio de la vida laboral, con una cantidad de información para procesar y un resultado concreto al cual llegar, pero que admite más de una solución. Es una tarea compleja, bien definida, desafiante para el alumno y realista. Coloca a los estudiantes ante una situación plausible, con variedad de datos a procesar y con cuestiones a resolver que requieren de la comprensión y uso de un conjunto de saberes geométricos y matemáticos. Tiene un propósito claro -elaborar un informe escrito- y un destinatario -el jefe de compras de la tienda-. Esto implica que el estudiante asuma un rol y trabaje con una audiencia "realista" en mente, no para el maestro o profesor. La tarea incluye ciertas restricciones que deberá tener en cuenta al momento de proponer una solución: el largo y el ancho del papel que deberá usar para envolver las cajas, así como las características de la ropa y sus dimensiones (seguramente no es lo mismo envolver una chaqueta que un sombrero). La tarea no tiene una única solución, sino que podría resolverse por caminos diferentes. Por tanto, incluye cierto grado de incertidumbre y requiere del estudiante, ser creativo, idear soluciones, imaginar distintas posibilidades y tomar decisiones. A la vez, esto posibilita discutir luego las distintas soluciones -informes- elaboradas por los estudiantes y analizar cuáles son mejores. Finalmente, es importante notar que se explicita al estudiante con qué criterios va a ser valorada su producción.

En términos de las grandes categorías de procesos cognitivos analizados en el capítulo anterior, la actividad claramente involucra los más complejos. Los estudiantes deben valorar diversas alternativas (formas y tamaños de cajas, costos) y crear un informe con recomendaciones sustentadas matemáticamente.

Como es obvio, esta actividad de evaluación no puede ser implementada en un único período de clase, sino que requiere de varias sesiones de trabajo. Es una actividad que, si bien podría ser propuesta para trabajar en forma individual, en realidad parece más adecuada para el trabajo colectivo, en pequeños grupos de dos o tres estudiantes. En todo caso, no debería ser propuesta como actividad de evaluación si los estudiantes no han tenido antes oportunidades de realizar tareas similares. Destinarle varias sesiones le aporta realismo a la actividad, ya que en la vida real este tipo de cuestiones no se elaboran ni resuelven de una vez. Seguramente será necesario que los estudiantes realicen aproximaciones iniciales al problema, las discutan colectivamente y con el docente, y elaboren un primer borrador que pueda ser revisado antes de entregar la versión final.

Un segundo ejemplo, referido a las ciencias sociales, está tomado de un trabajo de Russ Allen (s/fecha) (Recuadro 3.2).

En este caso la autenticidad de la situación radica en que reproduce un modo de construir conocimiento propio de la disciplina en cuestión, y en que la actividad termina con una presentación y discusión "real" de la información producida. En términos de los procesos cognitivos analizados en el capítulo anterior, la actividad requiere crear (recoger evidencia a partir de las entrevistas, interpretarla, idear formas para presentarla a los compañeros) y valorar el conocimiento construido, a la luz del producido por quienes se dedican a la disciplina (historiadores). Parece bastante claro que la actividad es más desafiante, motivadora y auténtica que una más tradicional como, por ejemplo, estudiar la Gran Depresión en un libro y luego responder preguntas sobre el tema en una prueba escrita.

Recuadro 3.2: Los efectos de la Gran Depresión en la vida de las personas

La tarea se compone de seis pasos:

1. Como grupo de estudiantes deben desarrollar un conjunto común de procedimientos, preguntas para entrevistas y un cuestionario. Este cuestionario será utilizado por cada estudiante para entrevistar a un adulto que haya vivido durante la Gran Depresión (ver Paso 2). Las preguntas deben estar basadas en el tipo de información que los integrantes del grupo desean recoger. Algunos ejemplos: ¿Cómo afectó la Gran Depresión a tu familia? ¿A ti personalmente? ¿A otros que conozcas o de los que hayas oído hablar? ¿Cómo eran las oportunidades de trabajo? Pensando hacia atrás, ¿cómo influyeron en tus valores y comportamientos de adulto las experiencias vividas durante la Gran Depresión?
2. Cada estudiante identifica y entrevista a una persona que vivió durante la Gran Depresión. Durante la entrevista el estudiante debe tomar notas cuidadosamente. (Nota: los estudiantes pueden grabar la entrevista). Después de completar la entrevista cada estudiante debe realizar una síntesis escrita del proceso y las respuestas.
3. Una vez que todas las entrevistas han sido completadas, los estudiantes se dividen en pequeños grupos (de cuatro o cinco integrantes) y discuten (comparan y contrastan) sus experiencias con las entrevistas, con el fin de llegar a algunas conclusiones sobre lo que han aprendido.

4. Cada estudiante debe consultar relatos de al menos dos historiadores que hayan escrito sobre la Gran Depresión, con el fin de determinar en qué medida las descripciones de los historiadores coinciden con las conclusiones del grupo. Se deben identificar y explicar las similitudes y diferencias.
5. Cada grupo debe desarrollar una presentación oral (de alrededor de 30 minutos de duración), en la que cada estudiante tenga un rol. Se trata de un informe a la clase, en el que los estudiantes deberán utilizar videos, audios, diapositivas u otros medios. Luego de la presentación habrá un espacio para preguntas.
6. Cada estudiante deberá escribir un breve reporte (de 5 o 6 páginas de extensión) en el que sintetice lo aprendido a lo largo de toda la experiencia.

Fuente: Allen, R. (s/f).

Como seguramente advertirá el lector, el trabajo con situaciones auténticas es algo que involucra tanto a la evaluación como a la enseñanza. No tendría sentido usar las situaciones auténticas exclusivamente en el momento de la evaluación, en especial si la finalidad de la misma es establecer una calificación. Si se piensa, en cambio, en evaluación formativa, un proyecto de trabajo basado en el enfoque de situaciones auténticas puede ser, a la vez, una propuesta de enseñanza/aprendizaje que incluya instancias de evaluación. Al concepto y práctica de la evaluación formativa le dedicaremos el capítulo que sigue, en tanto que el capítulo cinco estará enfocado en las evaluaciones cuyo propósito es establecer calificaciones.

2. La persistencia de las evaluaciones descontextualizadas en América Latina

En contraposición con los ejemplos anteriores y tal como adelantamos en el capítulo anterior, las actividades de evaluación en el aula que relevamos en nuestras investigaciones en varios países de la región, mayoritariamente carecen de situaciones auténticas. Tanto en primaria como en secundaria básica, predominan en forma abrumadora las actividades sin contexto o que plantean situaciones puramente escolares. Es llamativa la forma en que, a pesar de la enorme cantidad de iniciativas renovadoras y reformas curriculares ocurridas durante las últimas décadas en América Latina -de las cuales

presentamos antes apenas tres ejemplos-, persiste con fuerza la presentación y evaluación de los contenidos de estudio en forma descontextualizada.

Figura 3.2: Actividad de evaluación en Matemática en 6º de Primaria

4. 7,055 Dm: 7055 cm 0,00000000000007055
5. 0,005 m: 0,005 mm 0,000005 Km
6. 1,27 Gm: 0,00000000 127 Em
1270 000000 m.
(Repaso) Unidades de masa.
1 750 Mg: 7500000 Dg. 750000000 cg.
2. 10 804 Kg: 0,0000010804 Tg
10804000000 mg.
3. 3,58 Gm: 3580000 Kg
3580000000 g.
4. 86 dg: 8600000000 Ng. 0,000000000000086 Tg
5. 2300.000 cg: 23000000000 Mg 23 Kg
23000 g

Fuente: Base de registros del estudio **La evaluación en primaria en ocho países de América Latina.**

En la Figura 3.2 presentamos una típica actividad escolar de evaluación en Matemática en 6º de Primaria, similar a la mayoría de las presentadas en el capítulo anterior. Se trata de una tarea enfocada en conocimientos aislados, sin situación ni contexto y especialmente alejada de la vida real. El propósito del docente es evaluar la capacidad de los estudiantes para convertir cantidades en distintas unidades de medida. Sin embargo, las unidades de medida que se presentan en la tarea no suelen ser utilizadas ni en la vida cotidiana, ni en el trabajo científico. Hoy nadie escribe estas cantidades, por lo menos de esta manera. En el trabajo científico las cantidades extremadamente grandes o pequeñas se expresan como potencias de 10. Por ejemplo, para el ejercicio 5 de la primera parte de la Figura 3.2, la cantidad 0,000005 km se expresa en el ámbito científico como 5×10^{-6}. Por otro lado, la mayoría de las unidades de medida utilizadas (decigramos, centigramos, teragramos) tampoco son usuales en la vida corriente ni en la actividad científica (al igual

que los decilitros, centilitros y hectolitros). Tal vez los únicos lugares donde se encuentran estas unidades de medida en la actualidad sean la escuela y los textos escolares. En cambio, hay otras unidades de medida que hoy son de uso cotidiano, como los kilobytes, megabytes y gigabytes, que permitirían trabajan en torno a las mismas habilidades con mayor sentido. Por ejemplo, se podría proponer a los estudiantes que calculen en forma aproximada la cantidad de canciones que podrían almacenar en un disco de 8 gigabytes.

Figura 3.3: Actividad de evaluación en Química en 3º de Secundaria básica

EXAMEN DE QUÍMICA				TERCER AÑO		
1 - Completa la siguiente tabla:						
Elemento	**Símbolo**	**Z**	**A**	**Nº protones**	**Nº neutrones**	**Nº electrones**
Nitrógeno			14			
		13	27			
	Fe		56			
				11	12	
			31			15

2 -

a) Explica en qué consiste una DESTILACIÓN.
b) Explica en qué consiste una DECANTACIÓN.
c) Indica cómo separarías los componentes de un sistema formado por piedras, limaduras de hierro y cloruro de sodio. ¿Se trata de un sistema homogéneo o heterogéneo? ¿Por qué?

Fuente: Base de registros del estudio **La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina.**

En el mismo sentido, la Figura 3.3 ilustra con crudeza el tipo de aproximación al conocimiento científico que predomina en las propuestas de enseñanza y de evaluación. En la primera pregunta el estudiante debe conocer de memoria una gran cantidad de aspectos de los elementos químicos. La pregunta no plantea ninguna situación, simplemente es necesario haber memorizado todos los elementos, sus símbolos y composición. En la segunda pregunta el estudiante debe "explicar" en forma completamente descontextualizada el significado de ciertos términos utilizados en la disciplina para designar procesos vinculados a las soluciones. El ítem c) de esta pregunta es particularmente llamativo, por el "sistema" propuesto: piedras, limadura de hierro y cloruro de sodio. Con total franqueza, no sabemos si existe en la rea-

lidad un sistema formado por estos componentes. Si así fuese, lo lógico sería haber planteado una situación real en la que dicho sistema tuviese sentido.

En estas actividades, al igual que en buena parte de las presentadas en el capítulo anterior, no es posible encontrar prácticamente ninguna de las características de una situación auténtica: las tareas se presentan por fuera de cualquier contexto y no requieren del estudiante creatividad ni variedad de estrategias. ¿Qué le aporta a los estudiantes realizar estas actividades? ¿Qué tan motivados pueden sentirse ante este tipo de actividades? Estas son algunas de las preguntas que los docentes deberíamos hacernos al momento de elaborar nuestras propuestas de enseñanza y de evaluación. Si bien es cierto que no siempre es fácil encontrar contextos reales para las temas que trabajamos en el aula, deberíamos hacer un esfuerzo por acercar la vida escolar a la vida real y construir situaciones más cercanas a la realidad. Pensemos que el fracaso en responder este tipo de tareas es lo que determina que una parte importante de los estudiantes reprueben sus cursos y no completen la educación obligatoria que, por otra parte, hemos definido como un derecho.

3. Caracterización de las actividades auténticas

Las actividades auténticas como herramienta para evaluar los aprendizajes son realistas, tienen un propósito genuino con valor más allá del aula y una audiencia definida. Presentan cierto grado de complejidad. Muchas veces requieren ser trabajadas en colaboración entre pares, utilizando diversidad de recursos. Implican un trabajo en varias fases o etapas para llegar, de manera progresiva, a uno o varios productos previamente definidos y orientados por el docente a través de instrucciones y modelos de acción. A través de ellas los estudiantes evidencian su comprensión del contenido trabajado y su capacidad para utilizarlo en variadas situaciones propias de la vida en sociedad. A continuación analizaremos en forma particular cada una de estas características.

- Son realistas y plausibles y emulan del modo más cercano posible a la realidad. Las actividades auténticas tienden un puente entre la escuela y la vida fuera de ella. De esta manera, las actividades de evaluación adquieren significado porque son percibidas como relevantes para la vida. Si el fin último de la educación es garantizar la formación de ciudadanos críticos, reflexivos y competentes en distintos ámbitos -individuales, sociales, familiares, profesionales, científicos y laborales-, entonces las propuestas de evaluación en el aula deben incorporar la valoración de esas características en el vínculo de los estudiantes con el contenido enseñado.

- Son complejas e intelectualmente desafiantes. Generalmente proponen preguntas o plantean situaciones a las que los estudiantes deben dar respuesta a través de un trabajo de creación, valoración y/o investigación. No alcanza con recordar y aplicar conceptos o procedimientos. Muchas veces requieren de algún tipo de colaboración y trabajo en equipo. No evalúan conocimientos aislados, atomizados, arbitrarios ni artificiales. Son tareas relevantes y cargadas de significado para el estudiante, por lo que resultan interesantes y motivadoras.
- Tienen un propósito definido, una finalidad que requiere buscar soluciones a situaciones novedosas. Generalmente se propone alcanzar un producto que realmente alguien necesita o puede necesitar en alguna situación concreta (real o simulada). Por ejemplo: elaborar una guía turística, escribir un artículo para un periódico, organizar una presentación pública en torno a un tema, escribir un cuento o poema que será publicado para los padres, realizar una encuesta en la comunidad, escribir una solicitud de empleo, organizar un evento, elaborar recomendaciones para intentar resolver una situación problemática en la comunidad barrial, diseñar un experimento científico que permita corroborar o rebatir ciertas conclusiones, entre otras.
- El producto final va dirigido a un público determinado, a destinatarios o audiencias reales, más allá del docente. Son productos para ser presentados a otros: a una institución social, a una repartición de gobierno, a una autoridad escolar, a un grupo de padres o a la opinión pública, por mencionar algunas posibilidades.
- Ponen al estudiante en situación de desempeñar determinados roles, similares a los que llevan a cabo las personas en la vida real: se les propone asumir el rol de periodista, constructor, guía turístico, empleado de una tienda, moderador de un debate, científico, historiador o legislador, entre otros ejemplos posibles.
- Por el anclaje necesario que deben tener con las actividades propias de la vida real, los contextos de las tareas incluyen restricciones e incertidumbres. Los problemas que se proponen incluyen dificultades, limitaciones o aspectos restrictivos ante los cuales los estudiantes deben poner en práctica su creatividad, conocimientos y habilidades para sortearlos. Como las actividades admiten más de un camino o forma de realización, los estudiantes deben valorar alternativas y tomar decisiones. La creatividad muchas veces surge como forma de enfrentar y superar las limitaciones de la realidad.

- Requieren que los estudiantes pongan en juego un variado repertorio de estrategias: investigar, ensayar distintas soluciones, consultar variedad de recursos, practicar, probar distintas alternativas, realizar ajustes y volver sobre sus producciones preliminares para refinarlas. A lo largo del proceso son importantes las devoluciones del docente para orientar y reorientar la tarea.
- Debido a la complejidad que generalmente tienen, están pensadas para que los estudiantes las desarrollen durante un período de tiempo más o menos extenso, no de una sola vez. Suponen el desarrollo de un proceso de trabajo que se puede dividir en fases a las que se les asigna tiempos realistas de ejecución y que terminan con la elaboración de un producto final o la resolución de un problema (Perkins, D. 2010).
- Generalmente se desarrollan en contextos colaborativos, tal como ocurre en la vida real. Por lo general las personas, científicos, trabajadores y profesionales, realizan su trabajo en colaboración e intercambio con otros. Esto no significa que siempre que se propongan actividades auténticas se deba trabajar en equipo. Existen situaciones en la vida real que demandan el trabajo en forma individual y de esta manera deberían ser propuestas a los alumnos.
- La autoevaluación y la coevaluación son aspectos esenciales de la evaluación auténtica. Están siempre presentes en estas propuestas por ser procesos propios del trabajo individual y colectivo en situaciones reales. Mientras la autoevaluación favorece los procesos de autorregulación del aprendizaje, la coevaluación propende a desarrollar un papel activo y de ayuda recíproca donde se reciben aportes y se contribuye con nuevas ideas (sobre este aspecto nos extendemos con más detalle en el siguiente capítulo).

La Figura 3.4 sistematiza las diferencias existentes entre las tareas típicamente escolares, enciclopedistas, focalizadas en la adquisición del conocimiento en sí mismo y en el trabajo individual, y las tareas auténticas que promueven procesos de trabajo, el desarrollo de habilidades, el uso del conocimiento para resolver problemas con significado real y el trabajo colaborativo.

Figura 3.4: Diferencias clave entre las pruebas típicas y las tareas auténticas

PRUEBAS TÍPICAS O HABITUALES	TAREAS AUTÉNTICAS	INDICADOR DE AUTENTICIDAD
Esperan o requieren una única respuesta correcta.	Requieren un producto o desempeño de calidad y una justificación de las soluciones encontradas.	Se evalúa si el estudiante puede explicar, aplicar, realizar ajustes o justificar sus respuestas, y no solo si son correctas y utilizan los datos y algoritmos adecuados.
Las pruebas no pueden ser conocidas con antelación por los alumnos para garantizar su validez.	El producto o desempeño esperado es conocido con la mayor antelación posible; involucra excelencia en tareas exigentes pero predecibles.	Las pautas y criterios con que se juzgará el trabajo son predecibles y conocidas.
Carecen de un contexto plausible y de restricciones o limitaciones realistas.	Exigen el uso del conocimiento en contextos propios del mundo real: los estudiantes deben "hacer" historia, ciencia, matemática en situaciones simuladas pero realistas.	La tarea incluye un desafío y un conjunto de restricciones que son auténticas -como las que pueden encontrar en la vida real los profesionales, los ciudadanos, los científicos o los consumidores-.
Proponen preguntas aisladas que requieren el uso o reconocimiento de respuestas, información o procedimientos conocidos.	Constituyen desafíos integrales en los que el conocimiento y la valoración deben ser usados de manera innovadora para llegar a un producto de calidad.	La tarea es multifacética y no rutinaria, aún cuando haya una respuesta "correcta". Requiere clarificar o definir el problema, realizar intentos, equivocarse, realizar ajustes y adaptarse a una nueva situación.
Las tareas son simples, de modo que puedan ser calificadas en forma fácil y confiable.	Involucran tareas, criterios y estándares complejos y no arbitrarios.	La tarea involucra aspectos y/o desafíos fundamentales en el campo de estudio. No son fáciles de calificar, pero no se sacrifica la validez por la confiabilidad.
Se llevan a cabo de una sola vez. Los estudiantes tienen una sola oportunidad para demostrar su aprendizaje.	Son iterativas: permiten revisiones, modificaciones y refinamientos.	El trabajo está diseñado para mostrar si el estudiante ha logrado un dominio y comprensión real del conocimiento y no solo una cierta familiaridad pasajera.
Sus puntajes están basados en correlaciones estadísticas sofisticadas.	Proporcionan evidencia directa, mediante tareas que han sido validadas en función de roles adultos y desafíos basados en el trabajo en las disciplinas.	La tarea es válida a primera vista. Por tanto, atrae el interés y el esfuerzo del estudiante, y es vista como adecuada y desafiante por estudiantes y docentes.

PRUEBAS TÍPICAS O HABITUALES	TAREAS AUTÉNTICAS	INDICADOR DE AUTENTICIDAD
El estudiante recibe un puntaje.	El estudiante recibe retroalimentación útil que le permite confirmar el resultado de su trabajo o realizar los ajustes necesarios.	**La prueba está pensada no solo para valorar el rendimiento, sino también para mejorar el desempeño futuro. El estudiante es concebido como el principal usuario de la información.**

Fuente: Adaptado de Wiggins (1998), p. 23.

4. ¿Por qué es importante evaluar con situaciones auténticas?

4.1. Motivación hacia el aprendizaje

Evaluar a través de actividades que plantean situaciones auténticas hace a la evaluación más interesante, tanto para el alumno como para el docente. Al respecto Menéndez (1999) sostiene:

"Es imprescindible atender al factor motivacional que potencia el desempeño en la resolución de la situación planteada. Las situaciones que no impliquen atractivo, que no comprometan social ni cognitivamente, pretendiendo un valor extrínseco inherente a la evaluación, no pondrán en juego la competencia real del alumno y, por lo tanto, los resultados no reflejarán los logros posibles..."

Al trabajar con situaciones auténticas, los estudiantes se convierten en parte activa del proceso de generación de conocimientos. Se les da la oportunidad de conectarse con situaciones que podrían enfrentar en distintos contextos sociales con un propósito: investigar en ciencias, escribir para informar, persuadir o entretener a otras personas, interpretar acontecimientos y documentos históricos, aplicar procedimientos matemáticos para resolver situaciones del mundo real, criticar, debatir sobre diversos temas, construir conocimiento con sentido y aprender para la vida.

César Coll (1988) se refiere a la necesidad de que *"los aprendizajes que los estudiantes lleven a cabo sean, en cada momento de la escolaridad, lo más significativos posible"*. Este autor plantea la importancia de que el alumno tenga una actitud favorable para el aprendizaje y sostiene que ello dependerá de que perciba el aprendizaje como lógico y de que pueda relacionar los nuevos conocimientos con los que ya conoce. Además, señala este autor, *"un aprendizaje realizado de forma significativa es, al mismo tiempo, un aprendizaje que*

tiene un elevado valor funcional, es decir, un aprendizaje útil, un aprendizaje que puede ser utilizado con relativa facilidad para generar nuevos significados".

La significatividad de las propuestas de enseñanza y de evaluación es una contribución esencial para la motivación de los estudiantes, lo cual constituye una cuestión clave en un momento en que la falta de motivación y la falta de sentido de lo que se intenta aprender está directamente relacionada con la deserción y la desafiliación de muchos adolescentes del sistema educativo. Si bien no está exclusivamente en manos del sistema retener a los alumnos, ya que existen causas individuales, sociales y económicas que atentan contra ello, también existen factores relacionados con lo que sucede dentro de las aulas. Las propuestas educativas de los docentes y la orientación pedagógica de los cursos, en tanto inciden directamente sobre la motivación de los estudiantes, se convierten en elementos importantes para propiciar la permanencia. Este es pues un motivo adicional para preocuparnos por diseñar actividades que despierten el interés de los alumnos, a las que les encuentren sentido y que puedan resolver en ámbitos de trabajo colaborativo.

En este punto es importante considerar también que en las últimas décadas ha habido una revolución en relación al uso de la tecnología, especialmente en lo que refiere a la comunicación y al acceso a la información. Internet nos permite entrar al mundo de la comunicación y la información de una manera fácil, atractiva y atrapante. La educación formal debe "competir", de alguna manera, con esos medios para mantener el interés y la atención de los estudiantes, especialmente en la educación media. Si nuestras propuestas de enseñanza no acompañan estos cambios y continuamos enfocados en actividades poco desafiantes, carentes de sentido y alejadas de la vida de los niños y jóvenes (como las presentadas en las Figuras 3.2 y 3.3), captar su interés y atención será muy difícil. Muchos estudiantes optarán por abandonar el sistema educativo, en tanto que otros, aun cuando permanezcan, no estarán aprovechando su experiencia educativa para desarrollar sus potencialidades. En cambio, cuando las actividades de enseñanza y de evaluación involucran contextos reales y desafiantes, cuando se relacionan con las experiencias e intereses de los estudiantes, cuando son atractivas por lo que proponen, logran "seducirlos" y comprometerlos con la resolución de la tarea, más allá del hecho de estar siendo evaluados.

4.2. Transferencia de conocimientos

Otro de los aportes de la evaluación auténtica es el de contribuir a desarrollar la capacidad de los estudiantes para transferir los conocimientos desarrollados en una determinada situación para resolver otras. Al respecto Monereo

(2009) afirma que *"la evaluación auténtica tiene un obvio poder de predicción: la resolución de un problema "auténtico" en clase, predice de forma contundente la resolución de un problema similar en un contexto no académico"*.

Cuando se trabaja con actividades auténticas, ya sean reales o simuladas, se promueve la capacidad de los estudiantes para usar el conocimiento más allá de las situaciones específicas que les hayamos planteado. Tal como afirma Gardner en los textos que analizamos antes, en la medida en que el estudiante aprende la forma de pensar propia de una disciplina, se apropia de un repertorio de alternativas y de estrategias que podrá transferir a una amplia gama de situaciones.

4.3. Actitudes ante las instancias de evaluación

Otro aspecto relevante de este enfoque es el relativo a la ansiedad que las instancias de evaluación generan en muchos estudiantes. La evaluación mediante actividades auténticas implica explicitar claramente a los estudiantes y con suficiente antelación, la tarea que deberán realizar y los saberes y capacidades que se espera que demuestren. Al respecto Wiggins (1998) afirma que cuando se trabaja con el enfoque de la evaluación auténtica, ya no se oyen en los salones de clase preguntas tales como: ¿es esto lo que va a estar en la prueba?; ¿qué va a poner el profesor en el examen? Como las metas de aprendizaje y los desempeños a demostrar en la evaluación están definidas en forma clara desde el comienzo, los estudiantes trabajan en una zona de mayor seguridad y menor tensión. Profundizaremos en estos aspectos en los capítulos que siguen.

Según afirmamos antes, trabajar en un marco de evaluación auténtica nos permite -y nos exige- integrar la evaluación y la enseñanza, en lugar de concebir a la primera como el punto final del proceso. Nos ayuda a centrarnos en lo que es importante que los estudiantes aprendan, así como a observar dicho aprendizaje en términos de desempeños y no en términos de puntajes y promedios. El desempeño del estudiante en tareas auténticas hace visible para el docente lo que ha aprendido, su dominio de los conocimientos trabajados y las habilidades que es capaz de poner en acción para resolver situaciones.

Al respecto cabe recordar que evaluación y enseñanza van de la mano y deben estar alineadas. No podemos enseñar de un modo y evaluar de otro. Evaluar a través de situaciones auténticas requiere utilizar este tipo de situaciones en la propuesta de enseñanza. Esto no significa que sea necesario enseñar exclusivamente a través de este tipo de actividades, pero sí que a lo

largo de un curso los estudiantes deberían tener la oportunidad de enfrentarse a varias de ellas.

Trabajar con situaciones auténticas incrementa las posibilidades de realizar evaluaciones formativas, ofreciendo a nuestros estudiantes ayuda ajustada a sus necesidades a lo largo del proceso de trabajo y creando instancias de devolución que les permitan confrontar por sí mismos lo que lograron, con lo que se esperaba que lograsen. Este tipo de devoluciones no son posibles cuando las tareas de evaluación se limitan a preguntas o ejercicios carentes de contexto o situación. Solo son posibles cuando la tarea a realizar tiene cierto grado de complejidad. Este tema, el de las devoluciones y la evaluación formativa, será trabajado en profundidad en el capítulo que sigue.

4.4. Las situaciones auténticas y el aprendizaje de un deporte

En su libro ya citado Wiggins (1998) propone que para comprender la importancia del trabajo con situaciones auténticas es útil pensar en cómo se aprende un deporte. Supongamos que queremos enseñar a jugar al fútbol a un grupo de niños y niñas a lo largo de un año. Comenzamos explicándoles el origen de este deporte y sus reglas. Luego estudiamos con ellos la historia de las competencias, los grandes equipos y los futbolistas que se destacaron a lo largo de su historia, tanto a nivel nacional como internacional. En algún momento les pedimos que escriban una monografía sobre ellos. Luego pasamos a la parte "práctica" de nuestro curso, que consiste en la realización de ejercicios específicos en una cancha: correr con la pelota, tirar penales, hacer pases, marcar a un jugador que avanza rápidamente, eludir obstáculos, etc. Cada dos meses hacemos una evaluación. Algunas evaluaciones son teóricas: los estudiantes deben responder preguntas sobre la historia del fútbol y sus reglas. Otras son prácticas: los estudiantes deben tirar tres penales, sacar un tiro de esquina y dar cinco pases a otros compañeros. Al final del año les damos una calificación que es el promedio de los puntajes obtenidos en dichas evaluaciones bimestrales. Y así termina nuestro curso, sin que hayan tenido la oportunidad de jugar un partido de fútbol completo y "verdadero". Cuando los estudiantes preguntan para qué sirve lo que les estamos enseñando, les respondemos que cuando sean grandes y puedan jugar un partido entenderán para qué les enseñamos todo esto.

¿Enseñaríamos a jugar al fútbol de esta manera? Probablemente no. Sin embargo, así es como solemos enseñar lenguaje, historia, matemática o ciencias, desde la escuela primaria hasta la universidad. Enseñamos discursivamente los fundamentos y la historia de cada disciplina, hablamos sobre sus principales contenidos, hacemos que los estudiantes lean sobre la misma, les

hacemos realizar actividades puntuales y aisladas (responder preguntas, buscar información, resolver ejercicios). En las disciplinas científicas realizamos algunas "prácticas" de laboratorio. Pero muy pocas veces, a lo largo de toda su vida escolar y liceal, los estudiantes se ven enfrentados al desafío de llevar adelante un desempeño completo y real, para una audiencia real, propio de cada disciplina. En otras palabras, nunca les hacemos jugar un partido de fútbol completo y en serio.

Trabajar con situaciones auténticas equivale a jugar un partido de fútbol. Jugar un partido es lo que le da sentido a todo lo demás.

5. Actividades auténticas y saberes disciplinares

En este punto conviene detenerse un momento para abordar una interpretación errónea que suele hacerse del enfoque de evaluación auténtica, que consiste en pensar que este postula un acercamiento exclusivamente práctico y utilitario al conocimiento, y que desconoce la importancia de los saberes propios de las disciplinas. Al inicio de este capítulo repasamos rápidamente la propuesta de Gardner, orientada al desarrollo en profundidad de la capacidad de pensar desde las distintas disciplinas y a la "comprensión". Comprender implica conocer teorías y aspectos conceptuales de la disciplina en cuestión.

No hay, por tanto, oposición entre la enseñanza de los conceptos y procedimientos fundamentales de una disciplina y el trabajo con situaciones auténticas. La verdadera oposición debería plantearse, en todo caso, entre la memorización y aplicación rutinaria de contenidos no comprendidos, por un lado, y el uso de conocimiento "comprendido" para resolver situaciones auténticas, por otro. Por ejemplo, cuando proponemos a los estudiantes una situación auténtica cuya resolución involucra el cálculo de un porcentaje, obviamente deberán disponer de información previa sobre ciertos conceptos (razones, fracciones) y rutinas algorítmicas (multiplicación, división), o deberán buscarlas y construirlas en ese momento con el fin de resolver la situación. Para dar respuesta a situaciones propias del mundo real, los estudiantes inevitablemente necesitarán conceptualizar, comprender y dominar ciertos conocimientos básicos de las diferentes disciplinas. Pero para comprender el sustento del conocimiento de una disciplina, los estudiantes deben tener cierto grado de motivación y, además, necesitan que el contenido sea significativo. Esto implica, según vimos en el capítulo anterior, que puedan vincularlo con los conocimientos, saberes y experiencias que ya poseen. Por esta razón es importante aproximarlos a la utilización práctica del contenido que pretendemos enseñar, en su mundo cercano, presente o futuro. Si están

cargados de sentido, los conocimientos que se adquieran serán más perdurables y transferibles.

Pero este problema requiere aún una explicación adicional. Tal como se explicó en el capítulo anterior, la situaciones auténticas no se limitan a situaciones de la vida cotidiana o social, sino que incluyen también las situaciones propias de la forma de construir el conocimiento en cada disciplina. El problema en este punto es la diferencia entre situaciones disciplinares puramente escolares y situaciones disciplinares auténticas, es decir, las que emulan el modo en que realmente se trabaja en cada disciplina. Ejemplos de situaciones típicamente escolares son: i) la presentada antes en este capítulo, en la Figura 3.2 (la conversión de unidades); ii) calcular la velocidad de un gato que persigue a un ratón en una trayectoria rectilínea (presentada en el capítulo anterior, en la Figura 2.8); iii) solicitar a los estudiantes que realicen en forma manual operaciones con raíces cuadradas, para las que hoy cualquier científico o matemático recurre a una calculadora.

Es posible, sin embargo, trabajar con situaciones auténticas "intradisciplinares". El propio Grant Wiggins, respondiendo a ciertas críticas, enfatiza que la evaluación auténtica no implica eliminar el trabajo con problemas matemáticos "puros" en esta disciplina[10]. Lo que importa es que se trate de problemas auténticos, es decir, similares o representativos de los que se plantean quienes trabajan en la disciplina. Esto trae aparejado, por definición, que se tratará de problemas abiertos y poco estructurados (por oposición a las tareas escolares que admiten un único procedimiento ya conocido y una única respuesta correcta). En la respuesta referida Wiggins propone su ejemplo favorito de una situación auténtica de matemática "pura", que involucra el Teorema de Pitágoras:

> *"Todos sabemos que* [en un triángulo rectángulo, en el que A y B son los lados perpendiculares entre sí y C es el restante] $A^2 + B^2 = C^2$. *Piensen un minuto en el significado literal de esta propiedad: el área del cuadrado construido sobre el lado A sumado al área del cuadrado construido sobre el lado B, es igual al área del cuadrado construido sobre el lado C. La cuestión a analizar es la siguiente: ¿necesariamente la figura a construir en cada lado del triángulo debe ser un cuadrado? ¿Podría ser verdadera una versión más general del teorema? Por ejemplo: ¿Es verdadero o no que el área de un rombo construido sobre el lado A más el área de un rombo construido sobre el lado B es igual al área de un rombo construido sobre el lado C? Experimenten con estas y otras figuras.*

---

10. https://grantwiggins.wordpress.com/2014/01/26/authenticity-in-assessment-re-defined-and-explained/. Acceso el 24/06/16.

A partir de los experimentos, ¿es posible proponer una versión más general de este teorema?"

Este ejemplo muestra qué es "hacer" matemática de manera auténtica, desde la perspectiva del autor. Esta argumentación aplica a todas las disciplinas. El foco de la evaluación auténtica no está puesto exclusivamente en situaciones propias de la vida cotidiana o social, sino también en situaciones similares a las que requiere el trabajo con el conocimiento en las disciplinas. La única condición es que las situaciones y problemas sean similares a las que realmente enfrentan (o han enfrentado en el pasado, en la historia de la disciplina) quienes se dedican a ella.

Volviendo un instante a la metáfora del fútbol, trabajar con situaciones auténticas no significa que el estudio de fundamentos conceptuales y la realización de ejercicios puntuales no sean importantes. Lo son. Pero solo adquieren sentido cuando se juega un partido real y completo, es decir, cuando se trabaja con situaciones auténticas.

6. Autenticidad en el contenido y autenticidad en las condiciones de realización de las tareas

Hasta aquí hemos enfatizado la necesidad e importancia de enseñar y evaluar utilizando, al menos en algunos momentos, actividades y situaciones auténticas. Sin embargo, construir este tipo de situaciones no está exento de dificultades, sobre todo porque el aula es un espacio institucional específico y separado de otras instancias de la vida cotidiana. Gulikers, Bastiaens & Kirschner (2004) sostienen que es importante distinguir cinco dimensiones para determinar la autenticidad de una actividad de evaluación: la tarea (¿qué hay que hacer?); el contexto físico (¿dónde hay que hacerlo?); el contexto social (¿con quién hay que hacerlo?); el resultado (¿cuál es el producto solicitado?); y los criterios de evaluación (¿cómo será valorado?). Veamos cada una de estas dimensiones.

6.1. La tarea

Una actividad auténtica propone a los estudiantes una tarea similar a la que realizan las personas en ámbitos reales y requiere que los estudiantes integren conocimientos, habilidades y actitudes como lo hacen las personas, los profesionales o los trabajadores en la vida real. En general se trata de tareas complejas, aunque el nivel de complejidad debería estar dado por la similitud con la actividad real que intentan simular. Dicho de otra manera, no de-

bería incrementarse artificialmente la complejidad de las tareas. Los autores plantean que, si bien la mayoría de los problemas auténticos son complejos, implican más de una disciplina y tienen múltiples soluciones posibles, también existen problemas propios de la vida real que son simples, estructurados y pueden requerir saberes de una sola disciplina. Lo importante es que los estudiantes perciban el vínculo de la actividad con una situación del mundo real y, en particular, que la relacionen con sus intereses personales.

6.2. El contexto físico

Es importante, aunque no siempre sea posible, que el contexto de realización de las actividades sea similar al de las actividades reales. Por lo general el contexto físico en el que se desarrollan las actividades que proponemos -el aula o la escuela- es más "confortable" que un contexto real. Piénsese, por ejemplo, en las diferencias entre un taller de mecánica automotriz real y el taller de una escuela técnica. En muchos casos, no disponemos de la diversidad de medios y equipamientos que existen en el mundo real. Por ejemplo, la diversidad de instrumentos de un laboratorio científico o de medios de comunicación en un centro de prensa. Por estas razones, a pesar de que las tareas sean auténticas, existe siempre cierta distancia con las condiciones físicas y materiales propias de la vida real (si no fuese así, no estaríamos en una escuela). En todo caso, lo importante es que los recursos y herramientas disponibles, las informaciones que se aporten y los tiempos que se otorguen para realizar la tarea, sean lo más similares que se pueda a los existentes en la vida real.

6.3. El contexto social

El contexto social es otro elemento importante a considerar con respecto a la autenticidad de las actividades de evaluación. Los autores sostienen que, en la vida real, trabajar en equipo a menudo es la regla y no la excepción. Es importante que los procesos sociales a través de los cuales se desarrollan las actividades de evaluación se parezcan a los procesos sociales propios de la situación real equivalente. Por ello, cuando la situación real exige colaboración, la evaluación también debería incluirla. Pero cuando en la situación real normalmente se trabaja en forma individual, la evaluación debería ser individual.

6.4. El producto de la evaluación

Otro elemento clave para la autenticidad de las actividades es que tengan como resultado un producto tangible y bien definido. En el marco pro-

puesto por estos autores, un producto auténtico debe poner de manifiesto el dominio de los saberes y capacidades que se quiere evaluar. Para ello debe ser producido de un modo tal que podría tener valor en la vida real, más allá del aula. En la medida de lo posible, debería ser presentado o defendido ante otras personas, ya sea en forma oral o escrita, como forma de demostrar que el dominio alcanzado es genuino. Este elemento es clave. Recuerde el lector los ejemplos de "la sección de envolver regalos" y de "la Gran Depresión". No es lo mismo realizar la tarea y relatarla, que plasmar el resultado de la misma en un informe escrito o en una presentación oral con apoyos audiovisuales y preguntas de la audiencia. La realización y refinamiento de este tipo de productos es parte esencial del aprendizaje.

6.5. Los criterios de evaluación

La evaluación auténtica requiere que los criterios de valoración sean explícitos y transparentes, ya que en la vida real las personas (artistas, científicos, deportistas, trabajadores, profesionales, etc.) por lo general saben con qué criterios serán juzgadas sus actuaciones. Es importante que los criterios y requisitos del producto que se solicite a los estudiantes se relacionen con los criterios empleados normalmente en ese tipo de situaciones en la vida real. Al mismo tiempo, dado que se trata de un espacio educativo, también es importante que estén acordes al grado educativo y a las edades de los estudiantes a los que se les propone la tarea.

En suma, para que una evaluación sea auténtica debe plantear una tarea similar a las que se realizan en el mundo real (más allá del ámbito escolar), ser motivante, significativa y relevante. Pero, además, debe respetar o reproducir de la manera más fiel posible las condiciones de aplicación o el modo de realización de esa tarea en la vida real: tiempos, recursos, trabajo individual o colaborativo. Por ejemplo, si proponemos como actividad la elaboración de un artículo periodístico, se deberían plantear formatos, recursos, tiempos y fases de ejecución realistas, para que el "modo de producción" también sea auténtico. La redacción de un artículo para una revista o periódico, no puede ser realizada de una sola vez, en un único período de clase (que generalmente no excede los 60 minutos), escrito en el cuaderno de clase, en forma solitaria y con escasos recursos como fuente de información, porque estas no son las condiciones propias de la tarea en la vida real. Si bien la redacción de un artículo periodístico puede ser considerada una tarea auténtica, su autenticidad dependerá de que las condiciones de elaboración que se propongan a los estudiantes sean similares a las reales.

7. La autenticidad: un problema de grados

Proponer actividades auténticas y plausibles en el marco de la labor escolar no es fácil. Como en todo lo que aspiremos a concretar en nuestras prácticas, deberíamos comenzar por analizarlo desde la perspectiva de nuestros alumnos. Corresponde entonces preguntarnos ¿qué es auténtico para nuestros alumnos? En palabras de Monereo (2009): *"Lo que es realista, relevante y socializante para los docentes, no tiene por qué serlo para los alumnos y resulta imprescindible compartir esos criterios con ellos para que el carácter de autenticidad de las actividades se mantenga"*. Es preciso dialogar con ellos para averiguar qué es lo que perciben como auténtico en su entorno social.

Es fundamental tener en cuenta que lo auténtico no lo es siempre para todos, o por lo menos no lo es de la misma manera y con la misma intensidad en unos grupos que en otros. Influyen en ello variables como el nivel educativo, la edad, los intereses personales o el conocimiento práctico que tengan los estudiantes a partir de sus experiencias en su entorno directo y en la sociedad (Nicaise, Gibney y Crane, 2000). Un ejemplo puede ayudar a ilustrar esta dificultad. Un profesor de matemática plantea a los estudiantes un conjunto de preguntas sobre numeración basadas en una situación "auténtica": el retiro de dinero en un cajero automático, evitando los costos derivados del exceso de operaciones de extracción. Pero se trata de estudiantes de 8º grado, que no tienen experiencia previa de operar en cajeros automáticos, por lo que no llegan a comprender las condiciones de la situación (obviamente, podría ser una oportunidad para ir a un cajero automático y experimentar cómo funciona). De todas formas lo anterior no implica que solo se deba plantear a los estudiantes situaciones familiares. Por el contrario, un rol central de la educación es crear para los estudiantes oportunidades para conocer nuevas realidades. Las situaciones auténticas novedosas o no conocidas son de fundamental importancia, con la condición de que resulten comprensibles y motivadoras, lo cual requiere mucha intuición y conocimiento de sus estudiantes por parte del docente.

Otra dificultad a considerar, que sin dudas constituye un reto, es que las condiciones de realización de las actividades sean razonablemente realistas. La mayoría de las veces no es factible situar los conocimientos y capacidades que queremos enseñar y evaluar en situaciones completamente auténticas. En general lo posible es construir situaciones simuladas que recreen de manera más o menos fidedigna las condiciones propias de la situación real, aunque dentro del aula. Podemos hablar entonces de diferentes grados de autenticidad. Las actividades y situaciones no son "auténticas" o "no auténticas" de manera absoluta, existen diferentes grados de autenticidad, según las tareas se acerquen en mayor o menor grado a las condiciones reales.

Wiggins (1998), en su obra ya referida, sintetiza y ejemplifica la cuestión de los grados de autenticidad de la manera planteada en la Figura 3.5.

Figura 3.5: Grados de autenticidad

Inauténtico	Más o menos realista	Auténtico	Real
Analizar o responder preguntas a partir de un conjunto de datos	Diseñar una casa utilizando fórmulas matemáticas y formas geométricas predefinidas	Diseñar y construir una maqueta de una casa que satisface ciertos estándares de construcción y pedidos de los usuarios	La actividad sería real o completamente auténtica si alguien fuese a construir realmente esa casa
Escribir un ensayo sobre las leyes	Escribir un ensayo persuasivo sobre por qué una determinada ley debería ser modificada	Escribir una propuesta para ser presentada a los legisladores adecuados para promover la modificación de una ley existente	La actividad sería real o completamente auténtica si la propuesta fuese efectivamente presentada a los legisladores
Leer un fragmento de texto seleccionado por el maestro	Leer para toda la clase un texto seleccionado por el estudiante y de su interés	Hacer la grabación de audio de una historia, que pueda ser usada como audiocuento en una biblioteca	La actividad sería real o completamente auténtica si el audiocuento fuese usado realmente en una biblioteca

Fuente: Traducido y adaptado de Wiggins (1998: 28). La cuarta columna no está en el original, fue agregada por los autores

Las actividades propuestas en el capítulo anterior (Figuras 2.3, 2.9, 2.14, 2.15, 2.16, 2.17 y 2.18) constituyen buenas aproximaciones a situaciones auténticas, si bien la mayoría no son demasiado complejas ni satisfacen todas las características planteadas a lo largo del presente capítulo. Sin embargo, todas ellas se acercan a la autenticidad en diferentes grados. No hay que perder de vista que, lo que finalmente importa, es que las actividades que proponemos para el aprendizaje tengan significado real más allá del aula, que los estudiantes les encuentren sentido, que las perciban como relevantes y que comprendan por qué es importante abordarlas.

8. Transformando las consignas: una prueba típica a una actividad auténtica

Una primera forma de trabajar en la construcción de actividades auténticas que no signifique "comenzar de cero" sino aprovechar lo que ya

estamos trabajando en el aula, consiste en revisar las pruebas y actividades de evaluación que estamos utilizando y pensar en cómo transformarlas en situaciones auténticas. Se trata, en definitiva, de incorporar contextos significativos y situaciones realistas a los mismos contenidos que solemos presentar a través de tareas típicamente escolares, con el propósito de transformarlas en actividades que tengan propósitos concretos, destinatarios posibles, etapas de trabajo y un claro producto final a alcanzar.

Por cierto, lo deseable es realizar esta actividad junto con otros colegas. Al mismo tiempo, como ya fue señalado, se trata de un cambio que necesitaremos introducir no solamente en las actividades de evaluación, sino que requerirá también la revisión y adaptación de nuestras propuestas de enseñanza.

Para ilustrar en forma concreta nuestra propuesta de trabajo, presentaremos dos casos reales de experiencias de transformación de propuestas de evaluación, realizadas por docentes. El primer caso corresponde a la educación media y el segundo a primaria. En ambos casos las docentes parten de analizar una prueba típica propia, utilizada por ellas en clase, que luego transforman en una actividad auténtica en torno al mismo tema de la propuesta original.

8.1. El colesterol

Esta actividad fue elaborada por dos profesoras de Química en el segundo ciclo de Educación Secundaria, en un liceo público de Montevideo. Parten de una prueba sobre el colesterol constituida por cinco preguntas de respuesta abierta (ver Figura 3.6), dirigidas a evaluar los conocimientos de los estudiantes sobre el tema indicado. Según es posible apreciar y teniendo en cuenta el análisis realizado en el capítulo anterior, se trata de preguntas que requieren de los estudiantes recordar los conceptos trabajados en clase y reproducirlos para la docente, en una situación escolar, en la que los estudiantes se limitan a repetir lo trabajado en el aula o lo leído en los libros de texto.

Figura 3.6: Un típico "escrito" de Química en secundaria en Uruguay

Responde las siguientes preguntas:
1) ¿Qué es el colesterol?
2) ¿Qué diferencia hay entre el colesterol "bueno" y el colesterol "malo"?
3) ¿Qué consecuencias tiene el colesterol sobre la salud humana?
4) ¿Qué factores pueden causar un alza en los niveles de colesterol?
5) ¿Qué hábitos de alimentación y ejercitación contribuyen a evitar el colesterol malo?

Fuente: Trabajo realizado por las profesoras Inés Pena y Roxana Morey (2009).

A través del trabajo en equipo en régimen de Taller, las profesoras desarrollaron la propuesta incluida en la Figura 3.7, que evalúa los mismos temas, pero en el contexto de una situación auténtica, relevante y desafiante.

Figura 3.7: El "escrito" transformado en una situación auténtica

Química - Bachillerato en Medicina - Colesterol

1. Lean el siguiente texto:

   Las enfermedades cardiovasculares son la primera causa de muerte entre los uruguayos de más de 45 años. La prevención parece ser, de acuerdo a todas las investigaciones médicas, una de las formas de disminuir su aparición en la población joven. Existen factores de riesgo que aumentan las probabilidades de padecer estas enfermedades. Uno de ellos es poseer un elevado nivel de colesterol en sangre. Mediante un análisis de sangre en el cual se miden indicadores específicos, se puede determinar si una persona tiene un nivel de colesterol que ponga en riesgo su salud cardiovascular. Para ello, los médicos utilizan los llamados valores de referencia que indican los límites aceptados para hombres y mujeres.

2. Supongan que la Intendencia Municipal de Montevideo invitó a alumnos de su liceo del Bachillerato de Medicina a participar en una jornada de prevención en el marco del **"Día Nacional de Lucha contra las Enfermedades Cardiovasculares"**. Para ello deberán elaborar un boletín informativo que será distribuido en puestos especiales donde las personas podrán realizarse gratuitamente análisis de sangre para determinar sus niveles de colesterol.

3. Las pautas para elaborar el boletín son las siguientes:
   - Debe tener una extensión máxima de 4 carillas, c/u de las cuales será la mitad de una hoja tamaño A4.
   - Para la primera página del boletín deberán elaborar un eslogan que resuma el objetivo de la jornada y que resulte atractivo para las personas a las cuales está destinado.
   - Debe contener información clara y concisa sobre qué es el colesterol, por qué se habla cotidianamente de "colesterol bueno" y "colesterol malo" (HDL y LDL respectivamente), cuáles son los niveles de referencia normales para hombres y mujeres, qué relación existe entre los niveles anormales de colesterol y el riesgo de padecer un infarto.
   - Debe presentar recomendaciones en cuanto a la dieta más adecuada para las personas que presentan niveles anormales de colesterol. En particular, cuál es el consumo de grasas necesario y el balance entre grasas saturadas e insaturadas.
   - Debe indicar qué tipo de ejercicio físico se recomienda y por qué.

4. Para elaborar el folleto pueden consultar diferentes fuentes de información. A continuación les sugerimos algunas:
   a. Lehninger – Tratado de Bioquímica
   b. Tórtora – Anagnostakos,N – Principios de Anatomía y Fisiología
   c. www.msp.gub.uy (Ministerio de Salud Pública)
   d. www.suc.org.uy (Sociedad Uruguaya de Cardiología)
   e. www.nutinfo.org.ar

Fuente: Trabajo realizado por las profesoras Inés Pena y Roxana Morey (2009).

La situación propuesta es una campaña de sensibilización sobre aspectos sanitarios organizada por el Municipio, con puestos móviles distribuidos en la ciudad. Este tipo de campañas suelen realizarse tres o cuatro veces al año en Montevideo, con relación a distintos temas: cáncer de mama, presión arterial, semana del corazón, etc. En esa situación, los estudiantes tienen que elaborar un producto concreto y relevante, para destinatarios concretos: un folleto informativo para los ciudadanos. Esto implica realizar un esfuerzo especial por explicar los conceptos estudiados a alguien que no "sabe" sobre el tema (recordemos que una de las evidencias de la comprensión es la capacidad para explicar un tema a otros).

De esta forma, un escrito convencional, en el que los estudiantes debían escribir para el profesor sobre algo que el profesor ya sabe, se convierte en una situación auténtica en la que los estudiantes deben hacer un esfuerzo para comunicar el conocimiento a un público "real" que no conoce el tema. El conocimiento que se quiere evaluar queda situado en un contexto socialmente relevante. Se plantea un propósito que le da sentido a la tarea, así como un producto concreto y auténtico mediante el cual demostrar su comprensión del contenido involucrado. La tarea a realizar tiene restricciones y características propias de este tipo de productos: una extensión limitada y predefinida, la necesidad de incorporar elementos visuales y una frase o eslogan. Se orienta claramente a los estudiantes sobre lo que se espera que hagan y se les propone algunas posibles fuentes de información.

8.2. Un viaje a La Paloma

Esta actividad fue realizada por una Maestra de 6º en una escuela pública primaria de Montevideo. Al igual que en el caso anterior, parte de una propuesta de prueba de Geografía, sobre el departamento de Rocha (un departamento ubicado en el sureste de Uruguay, con costa sobre el océano Atlántico y fronterizo con Brasil). Las preguntas indagan sobre distintos aspectos de la geografía del departamento: ubicación, recursos económicos, flora y fauna, zonas balnearias y áreas protegidas. Al igual que en el caso anterior, se trata de preguntas cuya respuesta requiere de los estudiantes recordar conceptos trabajados en clase y reproducirlos para la docente.

Figura 3.8: Una típica prueba escrita en primaria en Uruguay

> Respondan a las siguientes preguntas:
> ¿A cuántos kilómetros se encuentra Rocha de nuestro departamento?
> ¿Cuál es el principal recurso económico de Rocha?
> ¿Cuáles son las principales especies vegetales y animales del departamento?
> ¿Cuáles son sus principales balnearios?
> ¿Existen áreas protegidas en este departamento? Si existen, averigua cuáles son y dónde están ubicadas.
> Ubica el departamento de Rocha en el mapa de nuestro país.

Fuente: Trabajo realizado por la maestra Graciela Loureiro (1993).

También en este caso la propuesta es transformada en términos de la situación a afrontar y los productos de la misma (véase la Figura 3.9), manteniendo los contenidos que la docente pretendía trabajar con sus estudiantes. En este caso la situación elaborada es la de un viaje de paseo de fin de año al departamento objeto de estudio. La docente organiza el trabajo en tres grupos, cada uno de los cuales tendrá a su cargo actividades diferentes, pero que deberán interactuar entre sí intercambiando información. La actividad pasa a incluir contenidos de matemática, al incluir como parte del producto la elaboración de un presupuesto para el paseo. También incorpora una "audiencia" o "destinatario" externo y real, al introducir la elaboración de un folleto para los padres. La actividad se desarrollará a lo largo de dos semanas, si bien en la propuesta no se especifica cuántos períodos de tiempo de clase serán asignados a la misma a lo largo de esas semanas. La propuesta prevé la elaboración y revisión de borradores parciales, que constituyen instancias apropiadas para la evaluación formativa. También se indican en forma escueta los criterios que serán empleados en la valoración de los productos finales, así como del proceso de trabajo colaborativo en los grupos. Un posible punto débil de la propuesta es que, dado que cada grupo tendrá una tarea diferente (recabar información, elaborar el presupuesto y preparar el folleto), no será posible valorar que todos los estudiantes comprendan y dominen los conocimientos trabajados.

Figura 3.9: La prueba transformada en una situación auténtica

Estamos en 6to año y tenemos muchos deseos de realizar un viaje de fin de curso. Todos votamos a qué lugar querríamos ir y decidimos que el lugar será La Paloma, en el departamento de Rocha. Para planificar este viaje comenzaremos ya a trabajar!!!

Para organizarnos les propongo las siguientes tareas para realizar en grupos.

Grupo 1

Para disfrutar al máximo de este viaje, primeramente será necesario buscar información sobre el lugar que visitaremos. Investiguen sobre: ¿a cuántos kilómetros se encuentra Rocha de nuestro departamento?; ¿cuál es la ruta principal de acceso para llegar?; ¿cuál es el principal recurso económico de Rocha?; ¿cuáles son las principales especies vegetales y animales?; ¿cuáles son los principales balnearios del departamento?; ¿existen áreas protegida en Rocha?; si existen, averigüen cuáles son y dónde están ubicadas. Ubiquen el departamento de Rocha en el mapa de nuestro país y el balneario La Paloma en el mapa del departamento. Esta información será útil para planificar el viaje y organizar las actividades que podríamos realizar.

Grupo 2

Los que integren este grupo serán los encargados de calcular los gastos. Deben prever el dinero necesario para pagar el transporte, el alojamiento, la comida y algunos paseos. Para su trabajo les puede ser útil algunos datos recabados por el grupo encargado de buscar información sobre el departamento, por ejemplo cantidad de kilómetros a recorrer para calcular el gasto del transporte, la ubicación de puntos interesantes para visitar: balnearios, áreas protegidas, etc.

Grupo 3

Este grupo tendrá como tarea preparar un folleto para los padres a fin de informarlos del viaje. Seguramente a ellos les interesará saber a dónde iremos, en qué fecha, por cuántos días, dónde nos alojaremos. Ustedes también deberán trabajar en forma colaborativa con el primer equipo, ya que necesitarán de los datos que ellos recojan para incluir información en el folleto.

POSIBLES FUENTES DE INFORMACIÓN

Visiten las siguientes páginas web:

- https://es.wikipedia.org/wiki/Departamento_de_Rocha
- http://www.turismorocha.gub.uy/
- http://www.rocha.gub.uy/portal/index.php
- http://www.mapadeuruguay.org/mapa-de-rutas-de-uruguay/
- https://www.google.com.uy/maps

Además, busquen en la biblioteca de la escuela libros sobre esta zona de nuestro país para obtener más información.

LAS TAREAS A REALIZAR

1. Discutan la manera cómo se distribuirán para integrar los tres grupos de trabajo. Para ello piensen qué tareas le gustaría a cada uno realizar. Cada grupo debería estar integrado por aproximadamente seis o siete compañeros.

2. Organicen el trabajo teniendo en cuenta los distintos roles que deberán desempeñar:
   a. Investigadores y recopiladores: buscar información en la web y en la biblioteca de la escuela sobre el departamento de Rocha. Organizar y seleccionar la información encontrada.
   b. Redactores y diseñadores: redactar y diseñar el folleto informativo.
   c. Encargados de las finanzas: calcular el dinero necesario que se deberá reunir.
3. Exploren los posibles recursos y fuentes de información (investigadores). No olviden consultar también el pronóstico del tiempo para los días del paseo.
4. Seleccionen y organicen los materiales, teniendo en cuenta la confiabilidad de las fuentes de información (recopiladores).
5. Redacten cuidadosamente la información que contendrá el folleto (redactores) Diseñen el material en papel de aproximadamente 12 cm x 15 cm. Asegúrense de que la información más importante, esté bien visible: con letra grande y en colores (diseñadores).
6. Para realizar los cálculos de dinero, tengan en cuenta que una de las primeras cosas que deberán confirmar es la cantidad de compañeros que viajarán, de esto dependerán los gastos de transporte, alojamiento y comida. Prevean posibles alternativas para el caso de que no se llegue a reunir el dinero necesario, por ejemplo, disminuir la cantidad de días o de posibles paseos a realizar en la zona. Las actividades que proyecten realizar deberán estar en función del dinero recaudado (encargados de las finanzas).

ETAPAS Y TIEMPOS

Cada uno de los grupos tendrá un tiempo aproximado de dos semanas para realizar su trabajo. Tengan en cuenta que los integrantes del grupo 1 serán los primeros que deberán poner "manos a la obra" porque de la información que ellos recaben depende, en gran parte, el trabajo del grupo 2 y 3. Estos últimos dos grupos podrán comenzar a trabajar dos o tres días después que los compañeros del grupo 1.

EVALUACIÓN

Cada grupo podrá elaborar borradores de las tareas realizadas para mostrármelos a mí y a los compañeros de los otros grupos y recibir sugerencias para mejorar los trabajos finales. La tarea de cada grupo será evaluada en función del trabajo colaborativo que realicen, la completitud y calidad de la información que recaben, el contenido y diseño del folleto, el realismo y exactitud del presupuesto elaborado.

Fuente: Trabajo realizado por la maestra Graciela Loureiro (elaborado en 1993, revisado y actualizado en 2016).

9. Resumen y sugerencias para el trabajo en el centro educativo

A lo largo de este capítulo hemos intentado explicar la importancia de incorporar situaciones auténticas a las propuestas de enseñanza y evaluación. Estamos convencidos de que ello es imprescindible si uno se propone que sus estudiantes vayan más allá del conocimiento superficial y declarativo. El trabajo con situaciones auténticas propicia la reflexión y la comprensión,

además de que genera más motivación en los estudiantes y hace que el aprendizaje sea más significativo.

La característica central de las situaciones auténticas es su similitud con las formas en que se usa el conocimiento en la vida real. Esto no implica en modo alguno una visión "utilitarista" y tampoco desconocer la importancia del conocimiento disciplinar. Por el contrario, tal como explicamos antes, las situaciones auténticas pueden ser las propias de la disciplina. La condición que deben cumplir es que se asemejen a los modos de construir el conocimiento por parte de quienes se dedican a la misma. Simultáneamente, son auténticas todas aquellas situaciones que se asemejan al uso que las personas hacen de los saberes en la vida social.

Introducir este tipo de cambios en nuestras prácticas pedagógicas no es una tarea fácil. Para hacerlo es necesario disponer de tiempo, trabajar con otros y, sobre todo, superar formas de trabajo que hemos ido construyendo a lo largo de años y que nos permiten sobrellevar el difícil desafío cotidiano del aula. No es fácil abandonar ciertas rutinas que nos dan seguridad en el día a día e intentar desarrollar formas de trabajo nuevas pero a la vez inciertas. Probablemente el mejor argumento para invitarles a intentar un cambio es el convencimiento de que este enfoque de la enseñanza y la evaluación conlleva ventajas importantes tanto para nosotros como para nuestros estudiantes. En el mediano plazo puede hacer más motivador, disfrutable y enriquecedor el trabajo en el aula.

Monereo (2009) explica que la modalidad de evaluación auténtica debe ser concebida como un proceso de enriquecimiento progresivo de las prácticas cotidianas de evaluación, y que las mismas deben ser construidas paso a paso, dentro de la "zona de desarrollo docente". No se trata, por tanto, de romper de un día para otro con las prácticas de evaluación que estamos empleando, sino de ir introduciendo, paulatinamente, cambios y mejoras en ellas.

Por otra parte, no todos los temas incluidos en el currículo de una disciplina pueden evaluarse a través de situaciones auténticas, o no siempre tendremos el tiempo para hacerlo. En ocasiones, y en función precisamente de los objetivos, los contenidos y las características de los estudiantes, las formas habituales de evaluación pueden ser perfectamente apropiadas (Monereo, 2009). La incorporación de situaciones auténticas en la enseñanza y la evaluación no implica dejar de lado actividades orientadas a evocar conceptos o realizar ejercicios sistemáticos para consolidar ciertos conocimientos y procedimientos.

Sin embargo, incorporar progresivamente algunas actividades auténticas a nuestros cursos puede contribuir a motivar a los estudiantes. A través

de las situaciones auténticas adquirirán un nuevo sentido los contenidos y procedimientos que normalmente enseñamos y evaluamos de un modo descontextualizado, o a través de situaciones y ejercicios escolarizados y artificiales. De allí la importancia de incluir al menos algunas de estas situaciones a lo largo de nuestros cursos.

Dada la complejidad del diseño de actividades auténticas, puede resultar muy útil el trabajo colaborativo con otros colegas y, en lo posible, insertarlo dentro de un proyecto institucional. En todo proceso de innovación el trabajo en equipo facilita la creación de nuevas prácticas, enriquece, permite sumar ideas y acortar tiempos.

Una forma concreta de empezar puede ser la revisión y transformación colectiva de las pruebas y actividades que estamos utilizando actualmente, al estilo de los casos presentados en el apartado anterior.

Otro camino posible es trabajar con los colegas en la conformación de un "portafolio de situaciones auténticas" de distinto tipo, que vinculen los conocimientos curriculares con situaciones y contextos reales propios de la vida social o disciplinar. Un portafolio docente de este tipo permite "tener a mano" un conjunto de tareas a las cuales recurrir como docente en forma individual, o para proyectos conjuntos entre varias "materias" o varios grupos de una misma institución. Cada docente podrá aportar ideas y materiales con su impronta y creatividad. A partir de las experiencias de uso en el aula, los diversos materiales podrán ir siendo modificados, actualizados y mejorados. También se podrían incluir en el portafolio las reflexiones de cada docente luego de cada experiencia.

Sugerencias para el trabajo en el centro educativo

A lo largo del proceso de elaboración de situaciones auténticas puede ser útil realizarse las siguientes preguntas. Las hemos adaptado y ampliado a partir de Wiggins (1998).

- ¿Es clara para los estudiantes la formulación del propósito de la actividad?

Al diseñar una actividad auténtica es necesario explicitar un propósito claro, concreto y realista para los alumnos. Este propósito deberá estar presente durante todo el desarrollo de la actividad. La formulación del mismo debe incluir la explicitación de los criterios con los que serán evaluados los productos o desempeños de los estudiantes. Nótese que estamos hablando del propósito de la actividad para los estudiantes, no de nuestras intenciones educativas ni de los aprendizajes que queremos promover en ellos.

- ¿Cuáles son las audiencias o destinatarios concretos de la actividad?

Las actividades auténticas no se realizan solo para ser mostradas al docente, sino que están dirigidas a destinatarios reales. Por eso, al diseñar una tarea auténtica debemos considerar a qué público estará dirigida (por ejemplo, alumnos de otros grados, padres, autoridades de la institución educativa, la comunidad barrial, una institución pública). En la mayoría de los casos será necesario simular la actividad, en el sentido de que el producto no llegará realmente a las audiencias, en tanto que en otros, y en la medida de lo posible, sí lo hará.

- ¿Es la situación razonablemente realista, relevante y compleja?

Es importante corroborar que la tarea se ha situado en el marco de una situación con cierto grado de complejidad, similar a las situaciones de la vida personal, social o disciplinar; y que sea relevante, es decir, que el producto a alcanzar represente algo de interés para alguien más allá del aula como, por ejemplo, un boletín o un informe de prensa para comunicar noticias, una feria científica para divulgar conocimientos, un proyecto de mejora para la comunidad, un libro de cocina con recetas para promover una alimentación saludable, un editorial para fundamentar opiniones, una encuesta para recabar datos, un afiche o video que formen parte de una campaña informativa, entre otros.

- ¿Incluye la actividad diseñada opciones y restricciones realistas?

Es importante asegurarnos que las tareas que diseñamos incluyan opciones y restricciones realistas. En el mundo real, cualquier tarea que emprendamos requerirá que tomemos decisiones en función de lo que queremos lograr y de las condiciones que nos rodean. Esas decisiones también dependerán de limitaciones y restricciones, como por ejemplo, de los recursos económicos disponibles, de los plazos a cumplir o de la disponibilidad de información que no siempre está fácilmente accesible.

- ¿Qué recursos requiere la realización de la actividad y dónde pueden encontrarse?

Otro punto importante a considerar es la necesaria orientación que debemos dar a los alumnos acerca de posibles recursos a consultar para realizar la tarea propuesta. Simultáneamente, muchas veces parte del desafío será que encuentren la información por sí mismos. Es importante incluir orientaciones hacia recursos útiles, variados y accesibles para obtener la información necesaria: libros, periódicos, revistas, sitios web, diversos recursos en línea.

- ¿Están previstas de manera razonable las diferentes etapas y productos parciales que permitirán llegar al producto final?

En la medida de lo necesario y dependiendo de la envergadura de la actividad, es importante prever distintas etapas o fases de trabajo, con tiempos realistas, como proceso a través del cual se podrá llegar al producto final. Si el proyecto tiene cierta complejidad, no podrá ser realizado en poco tiempo. Tampoco puede ser tan extenso que se corra el riesgo de cansar a los estudiantes y perder su motivación. Las etapas y/o productos parciales constituyen instancias clave para la evaluación formativa, tanto de parte del docente como entre los propios estudiantes.

- ¿Están definidos criterios claros e instancias para la autoevaluación, así como para la realización de devoluciones por parte del docente y de autoajustes por parte del estudiante a lo largo del proceso?

Al definir fases, etapas o productos parciales, es importante prever instancias de revisión, devolución y ajuste a lo largo de todo el proceso. Dichas instancias deben promover la autoevaluación y la coevaluación, de modo que los estudiantes desarrollen la capacidad de autorregular su aprendizaje (sobre estos aspectos ahondaremos en el próximo capítulo).

- ¿Están claramente explicitados los propósitos formativos de la actividad, así como los criterios y normas con que serán evaluados los productos o desempeños?

En el momento de proponer la actividad a los estudiantes, es importante compartir con ellos los propósitos formativos o de aprendizaje involucrados en la tarea (que son distintos de los objetivos del producto). La actividad tiene ciertos propósitos a lograr en términos del producto a alcanzar, por un lado, y ciertos propósitos formativos (de aprendizaje para los estudiantes), por otro. Los criterios de evaluación de ambos tipos de propósitos deben ser clarificados a los estudiantes, de modo que ellos puedan controlar el proceso.

- ¿Qué tan auténtica y relevante puede resultar la actividad para los estudiantes?

Por último, cuando elaboramos una actividad auténtica debemos preguntarnos siempre si lo que para nosotros es auténtico y relevante, lo será también para los estudiantes. La autenticidad de la tarea debe ser percibida como tal por el docente en el momento de elaborarla pero, fundamentalmente, debe resultar auténtica para los estudiantes. Esto requiere cierto grado de conocimiento de sus intereses y experiencias.

Referencias bibliográficas

Administración Nacional de Educación Pública (ANEP) (2002). *La Educación Media Superior Uruguaya en el Siglo XX*. Serie "Aportes para la reflexión y la transformación de la educación media superior", Cuaderno de Trabajo Nº 7. Montevideo: ANEP, Comisión de Transformación de la Educación Media Superior Pública.

Allen, R. (s/fecha). *Performance Assessment*. WEAC Instruction and Professional Development Division. Disponible en http://weac.org/articles/performance-assessment/. Acceso el 27/05/16.

Archbald, D.A. y Newmann, F.M. (1988). *Beyond standardized testing: Assessing authentic academic achievement in the secondary school*. Washington, DC: Office of Educational Research and Improvement.

Coll, C. (1988). *Significado y sentido en el aprendizaje escolar. Reflexiones en torno al concepto de aprendizaje significativo*. Barcelona: Universidad de Barcelona. Disponible en: http://dialnet.unirioja.es/servlet/articulo?codigo=48298. Acceso 05/01/17.

Ferreiro, E. y (compiladoras) (1982). *Nuevas perspectivas sobre los procesos de lectura y escritura*. Mexico D.F.: Siglo XXI Editores.

Ferreiro, E. (1986). *Proceso de alfabetización, la alfabetización en proceso*. Buenos Aires: Centro Editor de América Latina.

Ferreiro, E. y Terebosky, A. (1991). *Los sistemas de escritura en el desarrollo del niño*. Mexico D.F.: Siglo XXI Editores.

Ferreiro, E. (1998). *Alfabetización. Teoría y práctica*. Mexico D.F.: Siglo XXI Editores.

Frey, B. y Schmitt, V. (2012). *Defining Authentic Classroom Assessment*. En: Practical Assessment, Research and Evaluation. A peer-reviewed electronic journal, Volumen 17, Nº 2. Disponible en: http://pareonline.net/getvn.asp?v=17&n=2. Acceso el 27/05/16.

García Castro, A. (2014). *"Cada minuto de clase puede ser vivido como un acto de creación"*. Entrevista: Glauco Daniel Cabrera y el Plan Piloto de Educación de

Uruguay 1963. Publicado en la Revista Perspectiva (Revista Digital Docente), Nº 9, Junio 25, 2014. Disponible en: http://revistaperspectiva.cl/entrevista-glauco-daniel-cabrera/. Acceso 03/01/17.

Gardner, H. (1999). *The Disciplined Mind: What All Students should Understand.* New York: Simon & Schuster.

Gulikers, J. Bastiaens, T. y Kirschner, P. (2004). *A Five-Dimensional Framework for Authentic Assessment.* En: Educational Technology Research and Development, Vol. 52, Nº 3, pp. 67-86. Nueva York: Springer.

Instituto Nacional de Evaluación Educativa (INEEd) (2014). *Informe sobre el Estado de la Educación en Uruguay 2014.* Montevideo: INEEd.

Leymonié, J. y otros (2013). *La evaluación en las aulas de Ciencias Naturales en cuatro países de América Latina: Concepciones de Ciencia y de Educación Científica.* Documento no publicado que forma parte de los productos del estudio "La evaluación de los aprendizajes en Ciencias Naturales en las aulas de educación media en América Latina".

Macedo, B. (2006). *Habilidades para la vida: contribución desde la educación científica en el marco de la Década de la educación para el desarrollo sostenible.* Ponencia presentada en el Congreso Internacional de Didáctica de las Ciencias, La Habana, Cuba, febrero de 2006. Disponible en: http://unesdoc.unesco.org/images/0016/001621/162181S.pdf. Acceso 01/06/16.

Menéndez, I. (1999). *Evaluación.* En: Quehacer Educativo Nº 38, pp. 19-20.

Monereo, C. (2009). *La autenticidad de la evaluación.* En: Castelló, M. La evaluación auténtica en la enseñanza secundaria y universitaria: Investigación e Innovación. Barcelona: Edebé.

Nicaise, M., Gibney, T. y Crane M. (2000). *Toward an Understanding of Authentic Learning: Student Perceptions of an Authentic Classroom.* En: Journal of Science Education and Technology, Vol. 9, Nº 1, pp 79-94. Nueva York: Springer.

Organización para la Cooperación y el Desarrollo Económico (OCDE) (2006). *PISA 2006. Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemáticas y Lectura.* Madrid: Ministerio de Educación y Ciencia y Editorial Santillana.

Perkins, D. (2010). *El aprendizaje pleno. Principios de la enseñanza para transformar la educación.* Buenos Aires: Paidós.

Wiggins, G. (1989). *A True Test: Toward More Authentic and Equitable Assessment.* The Phi Delta Kappan, Vol. 70, Nº 9, pp. 703-713.

Wiggins, G. (1998). *Educative Assessment. Designing Assessments to Inform and Improve Student Performance.* San Francisco: Jossey-Bass.

CAPÍTULO 4
EVALUACIÓN FORMATIVA

Introducción

En los capítulos anteriores analizamos las actividades de evaluación que proponemos habitualmente a nuestros estudiantes. Nos enfocamos en los tipos de tareas y procesos cognitivos que las mismas requieren para su resolución, así como en el tipo de situaciones que involucran. Destacamos la necesidad de revisar y mejorar dichas actividades, de modo que su realización requiera de los estudiantes comprensión, reflexión, valoración y creación. Intentamos mostrar la importancia que para ello tiene el uso de situaciones "auténticas", es decir, similares a las que ocurren en la vida real, tanto en la vida en sociedad como en el trabajo propio de cada disciplina. Señalamos que este es un desafío que involucra tanto a la evaluación como a la enseñanza, dado que ambas deberían ir "de la mano" (no deberíamos tener un enfoque para enseñar y otro a la hora de evaluar). Por tanto, modificar la evaluación exige revisar también nuestra manera de enseñar. Ofrecimos pistas para construir situaciones en las que el uso del conocimiento tenga sentido más allá del ámbito del aula, para lo cual es importante que tengan un propósito, un producto y destinatarios bien definidos. Enfatizamos también la importancia de que las actividades puedan ser resueltas por diversos caminos y destacamos el potencial de las situaciones auténticas para involucrar y motivar a los estudiantes con su aprendizaje.

En este capítulo daremos un paso más en el análisis de las prácticas de evaluación, pero nos "despegaremos" de las actividades de evaluación formalizadas en pruebas, para enfocarnos en la evaluación formativa, en tanto elemento constitutivo clave en la articulación de los procesos de enseñar y de aprender. Analizaremos la evaluación como una herramienta que debería formar parte de la enseñanza cotidiana y que debería ser refinada y mejorada en forma continua.

Hasta el momento hemos trabajado en torno a las actividades de evaluación sin hacer referencia a su finalidad. En rigor, existen dos grandes formas o tipos de evaluación en el aula, que tienen finalidades y lógicas radical-

mente diferentes. Tradicionalmente se las denomina "evaluación formativa" y "evaluación sumativa". Estos términos fueron inicialmente acuñados por Michael Scriven en los años 60 con el fin de distinguir entre las evaluaciones que tienen lugar al final de un período de enseñanza y cuyo propósito es calificar al estudiante, y aquellas que tienen lugar durante el proceso de enseñanza y cuya finalidad es adaptar el proceso didáctico a los progresos y necesidades de aprendizaje observados en los alumnos, con el fin de contribuir a mejorar su desempeño (Scriven, M., 1967; Jorba, J. y San Mart, N., 1993).

Habitualmente los docentes realizamos ambos tipos de evaluaciones. Evaluamos para ajustar nuestra enseñanza y brindar orientaciones a nuestros estudiantes y, al mismo tiempo, evaluamos para calificarlos. Sobre la base de las calificaciones u otros registros, al final de cada año o semestre debemos tomar una decisión con relación a la aprobación o reprobación del curso por parte de cada estudiante. Para la mayoría de nosotros este tipo de evaluación es uno de los aspectos más ingratos de nuestro trabajo. Nos exige corregir gran cantidad de trabajos y, sobre todo, nos obliga a tomar decisiones difíciles, sobre la base de información limitada, que afectan la trayectoria educativa de los estudiantes. Nos coloca en una disyuntiva compleja. ¿Debería aprobar a este alumno aunque no haya logrado todo lo que yo espero de mi curso o lo que establece el programa de estudios? Si lo hago, ¿no estaré dando por satisfactorio un desempeño que no lo es y, en cierto modo, "bajando el nivel" educativo? Si lo repruebo, ¿no estaré obstaculizando el crecimiento del estudiante y su trayectoria educativa? ¿Es realmente importante que hayan aprendido esto que estoy evaluando, o es algo que la mayoría de las personas olvida rápidamente y no vuelve a necesitar a lo largo de su vida?

Con mayor o menor frecuencia la mayoría de los docentes nos hacemos este tipo de preguntas y preferiríamos no tener que calificar a los estudiantes. Sin embargo, hacerlo es parte esencial de nuestra responsabilidad profesional, al menos en el régimen académico que prevalece en la mayoría de los sistemas educativos: un sistema organizado en torno a grados y materias o áreas de conocimiento, que deben ser aprobados en forma semestral o anual para poder avanzar en los estudios[1]. En este marco, las familias, nuestros colegas, el resto del sistema educativo y la sociedad en general, esperan que demos cuenta del grado en que los estudiantes han aprendido los conocimientos, saberes y habilidades propias de cada curso que enseñamos. De esto depende que cuenten con los conocimientos necesarios para los cursos siguientes o para avanzar hacia el próximo nivel de enseñanza. A esta función social

1. El problema de las calificaciones y su función dentro del modelo educativo propio de la sociedad industrial será analizado en profundidad en el próximo capítulo.

de la evaluación la denominaremos "certificación". De acuerdo con la Real Academia Española, **certificar** significa *"asegurar, afirmar, dar por cierto algo"* y *"hacer constar por escrito una realidad de hecho por quien tenga fe pública o atribución para ello"*. En la educación, los docentes tenemos la atribución formal y la "fe pública" para asegurar o "dar por cierto" en qué medida cada estudiante ha logrado lo que se espera como aprendizaje en un curso. La nota o calificación es la "constancia escrita" y administrativa de este hecho. El sistema educativo y las familias depositan en nosotros su confianza para que desempeñemos esta función.

En el marco de esta conceptualización, más que de evaluación "sumativa", término que tiene una connotación negativa entre los docentes, preferimos hablar de *"evaluación para la certificación"*. E insistir en que se trata de un tipo de evaluación que tiene una lógica completamente diferente a la de la evaluación formativa, pero que es parte de nuestro rol profesional. La función de certificación es particularmente importante en la educación superior, dado que en este nivel la aprobación de los cursos tiene como consecuencia social principal la habilitación del estudiante para el desempeño de un rol, oficio o profesión (ejercer la medicina o la docencia o como electricista o analista de sistemas, entre otros). En todos estos casos las instituciones educativas y los docentes asumimos (o deberíamos asumir) una responsabilidad ante la sociedad por la preparación de los egresados para el desempeño adecuado de dichas funciones.

Existen pues dos tipos principales de evaluación de aprendizajes en el aula, con finalidades y lógicas propias. Sin embargo, como se mostrará en las páginas que siguen, en la práctica suelen aparecer entremezclados. La situación es bastante curiosa. En el discurso que tenemos los docentes, al igual que en buena parte de los discursos de quienes se dedican al estudio de la pedagogía y la didáctica, se suele enfatizar las bondades de la evaluación formativa. Es común que los docentes expresemos una marcada valoración de la función formativa de la evaluación. La percibimos como la "evaluación buena" y deseable, contraponiéndola a la evaluación sumativa, a la que en nuestro fuero interno visualizamos negativamente, como una imposición administrativa del sistema. Al mismo tiempo, la evaluación sumativa nos exige tiempo y esfuerzo, nos genera conflictos y nos resulta ingrata, según señalamos algunos párrafos más arriba. Y sin embargo, a pesar de todo esto, en las aulas predomina el uso constante de las notas y calificaciones por encima de la evaluación formativa, según veremos enseguida.

Este capítulo tiene como propósito aportar elementos para conceptualizar, revisar y mejorar las prácticas de evaluación formativa, tanto a través de las actividades de enseñanza y de aprendizaje, como en las instancias for-

males de evaluación. Comenzaremos por un breve repaso de los hallazgos realizados en los trabajos de investigación que llevamos adelante en varios países de la región, enfocándonos en las concepciones y prácticas de evaluación formativa que encontramos en maestros y profesores. En segundo lugar, discutiremos el concepto de evaluación formativa y sus implicancias para la enseñanza y el aprendizaje. En tercer lugar, propondremos un conjunto de estrategias e ideas concretas para mejorar la práctica de la evaluación formativa como elemento central en la enseñanza. Finalmente, analizaremos el potencial de las matrices de valoración o "rúbricas" para mejorar el papel formativo de la evaluación. En el capítulo siguiente nos enfocaremos en las evaluaciones con fines de certificación.

1. Concepciones y prácticas de evaluación formativa prevalecientes entre los docentes en América Latina

¿A qué nos referimos los docentes cuándo utilizamos el término "formativo" aplicado a la evaluación? ¿Tenemos una concepción compartida o visiones diferentes? ¿Estamos todos poniendo el énfasis en la función que cumple la evaluación en el proceso de aprendizaje de cada estudiante? ¿Qué tipos de devoluciones hacemos a nuestros estudiantes en las instancias formales de evaluación, para llevar a la práctica la función formativa?

Comencemos por la evidencia empírica que recogimos en los estudios que realizamos en los años 2008 y 2012, descritos en el capítulo uno. Dichos estudios incluyeron entrevistas en profundidad con los docentes en torno a la evaluación, así como un relevamiento y análisis de las actividades de evaluación propuestas por los docentes a sus estudiantes. Incluyeron también encuestas para recoger el punto de vista de estos últimos.

En el marco del trabajo realizado en 2012, que estuvo enfocado en profesores de ciencias naturales de educación media básica, se pudo identificar la existencia de tres formas principales de concebir la evaluación formativa por parte de los docentes (Ravela y otros, 2014: 35). Por un lado se constató, especialmente en Perú y Colombia[2], que los docentes asociaban el término "formativo" a la evaluación de los aspectos vinculados con la formación integral y a la formación en valores. Entendían la evaluación formativa como aquella que se enfoca en las actitudes de los estudiantes. En particular en Co-

2. Las referencias a lo característico de cada país está basado en la información relevada en los estudios mencionados, que no estuvieron basados en muestras representativas. Por tanto, no tienen la pretensión de realizar generalizaciones, sino de describir la evidencia relevada en nuestros estudios en cada país.

lombia, la mayoría de los entrevistados contrapuso la evaluación formativa con la evaluación de conocimientos de las disciplinas, como cosas diferentes. *"Su actitud en clase: su posición, cómo se debe sentar o caminar, y las normas sobre todo de convivencia... Por ejemplo, dejar el salón arreglado como lo encontramos, que las basuras no me las arrojen, que no se tiren objetos... se maneja mucho la parte de valores, la responsabilidad, por ahí le veo la parte formativa"*. Desde esta primera perspectiva, la evaluación formativa se vincula exclusivamente con la formación en actitudes y valores, pero no se la concibe como una herramienta que forma parte del proceso didáctico y que tiene una función en el desarrollo del aprendizaje en las distintas áreas del conocimiento.

Una segunda forma de concebir la evaluación formativa por parte de los docentes entrevistados la vincula con la constatación del avance de los alumnos hacia lo que se les pedirá en la evaluación sumativa. En palabras de un entrevistado, la evaluación formativa es *"la evaluación sin calificación"*, de lo que luego será calificado en una prueba. Esta concepción se constató especialmente en Chile, en que la enseñanza está fuertemente orientada hacia las instancias de calificación. *"Es un tipo de evaluación para que los estudiantes sepan qué les falta repasar antes de la prueba"*. En este caso, el foco se pone en preparar a los estudiantes para un mejor desempeño en la prueba de calificación o certificación, es decir, en la instancia de evaluación sumativa que, en cierto modo, pasa a ser la meta principal de la enseñanza.

Finalmente, un tercer conjunto de docentes conciben a la evaluación formativa como aquellas instancias que se enfocan en los procesos cognitivos a través de los cuales los estudiantes avanzan en sus aprendizajes. El foco de atención se pone en los procesos en sí mismos, más que en las instancias posteriores de evaluación sumativa. El testimonio que sigue ilustra esta perspectiva: *"estoy permanentemente preguntando qué entendiste por esto, no con ánimo de caerle con el mazo sino de ver si entendió, y si no entendió, de explicárselo o que otro se lo explique"*. Este tipo de preocupación estuvo presente sobre todo entre los docentes entrevistados en Uruguay.

En el transcurso de las entrevistas en profundidad, se preguntó específicamente a los docentes acerca del tipo de devolución que dan a sus alumnos luego de una prueba, con el fin de ayudarles a avanzar a partir de lo que ya lograron. Se constató que pocos docentes promueven una revisión reflexiva de sus trabajos por parte de los estudiantes. En general las devoluciones que realizan los docentes se enfocan en el señalamiento o corrección de los errores cometidos.

> *"En la codificación de las muestras de trabajos se registró qué tipo de intervención de los docentes había en los trabajos, cuando estos estaban respondidos por los alumnos. La evidencia es contundente en el sentido de*

que la devolución por escrito se limita a marcas de correcto/incorrecto o a la asignación de puntos y calificaciones. Apenas el 4% del total de las propuestas (es decir, considerando tanto las respondidas por alumnos como las no respondidas) de Colombia y el 3% en Uruguay, tenían algún tipo de comentario escrito del profesor. Estos comentarios, por otra parte, son bastante vagos: '¿Por qué no buscó otro video?', 'revísalo mejor', 'falta el dibujo', 'no clasificó cada enlace', 'falta unir los puntos (en una gráfica)', 'completa lo que falta'. Algunos (pocos) comentarios más específicos fueron: 'Bien, pero no aclaras cómo cuidar la piel, lee la consigna'. 'Sí pero ¿cómo? ¿Qué hay en la piel que nos permite percibir los estímulos como el cambio de temperatura o la vibración?'; 'no es esta la razón, tiene que ver con la parte a), la piel puede ser gruesa pero muy sensible'." (Ravela y otros, 2014: 32).

Esta situación fue corroborada por los propios estudiantes, a quienes se preguntó en un cuestionario qué hacían sus profesores cuando les devolvían una prueba escrita. De acuerdo a la información aportada por los estudiantes (Figura 4.1), lo que más frecuentemente hacen los docentes, es marcar los errores y las cosas que están mal. Lo menos frecuente en la experiencia de los estudiantes, es recibir comentarios explicando la diferencia entre lo que hicieron y lo que se esperaba que realizaran, así como tener la oportunidad de mejorar el trabajo realizado.

Figura 4.1: Porcentaje de estudiantes que informaron que frecuentemente o casi siempre sus profesores hacen lo indicado en la frase cuando les devuelven una evaluación escrita

	CHI	COL	PER	URY
1. Solamente te pone la nota en el trabajo	59,7	62,6	65,3	50,2
2. Te marca los errores y cosas que están mal	86,1	86,2	84,7	85,6
3. Te escribe un comentario explicando la diferencia entre lo que hiciste y lo que se esperaba que hicieras	39,4	35,3	46,1	36,7
4. Te dice lo que tienes que mejorar	55,2	59,2	71,6	49,8
5. Te dice que debes esforzarte y estudiar más	64,0	66,7	72,5	61,7
6. Te da la oportunidad de rehacer o mejorar tu trabajo	28,6	49,9	52,4	28,8

Fuente: Ravela y otros, 2014: 32.

En el estudio realizado en 2008 en 6º de primaria en ocho países, también se constató que son excepcionales los comentarios escritos de los docentes orientados a promover la reflexión de los estudiantes sobre lo realizado en una tarea de evaluación o a que la revisen en función de lo que se esperaba que hicieran. El estudio de las propuestas de evaluación escrita, seleccionadas

por los docentes visitados como ejemplos de sus prácticas cotidianas en el aula, mostró un claro predominio de la tendencia a entregar a los alumnos mensajes de estímulo (*'muy bien, te felicito'; tú puedes'*); exhortaciones a un mayor esfuerzo (*'necesitas esforzarte un poco más', 'tu tarea no ha estado completa', 'dedíquese', 'debe mejorar'*) y, además, puntajes y calificaciones. En varios miles de tareas revisadas en este estudio, prácticamente no se encontraron devoluciones descriptivas o reflexivas en los trabajos escritos de los estudiantes, sino casi exclusivamente marcas de acierto o error, calificaciones, puntajes, felicitaciones o indicaciones de que algo es correcto o incorrecto.

Como es natural, casi todos los docentes califican los trabajos de sus estudiantes utilizando la escala de notas usada en el país. Sin embargo, como en ninguno de los países en los que se realizó este estudio estaban claramente definidos los referentes conceptuales de dichas escalas[3] (por ejemplo, qué significa un desempeño suficiente en matemática), las notas ofrecen a los estudiantes solamente información acerca de su posición relativa con respecto al puntaje mínimo y al puntaje máximo y, en particular, de cómo está su trabajo con relación a los de sus compañeros. Casi no se registraron casos en que los docentes ofreciesen retroalimentación descriptiva, es decir, expresiones o comentarios que permitiesen al alumno comparar lo que había realizado con lo que se esperaba que lograra (Picaroni, B., 2009: 85-86).

En un tercer estudio, realizado entre 2012 y 2013 por un equipo de investigadores de la Universidad de Aguascalientes de México, se analizaron las "evidencias de prácticas de evaluación" de un grupo de docentes de 2°, 5° y 6° de primaria en escuelas de Nuevo León. A partir de dicho análisis los investigadores constataron que:

> *"La calificación de exámenes y tareas privilegia el conteo de errores, sin referencia a estándares de desempeño precisos... La retroalimentación se suele reducir a informar sobre el resultado de la calificación, en la forma de una nota a la que se llegó mediante un proceso que la hace carecer de significado pedagógico, que frecuentemente va acompañada de elogios o amonestaciones, así como de exhortaciones a mejorar, sin orientación sobre cómo conseguirlo"* (Mercado y Martínez Rizo 2014: 593).

Por tanto, también en este estudio se constató que el foco de los docentes estaba puesto en calificar en forma continua a los estudiantes sin referentes claros, así como el predominio de la práctica de señalar errores y la carencia de acciones dirigidas a promover el avance en los aprendizajes de

3. Analizaremos las escalas de calificación y los reglamentos de evaluación en el próximo capítulo.

los alumnos. En las conclusiones del estudio se indica, en primer lugar, que las tareas que se proponen a los estudiantes no son las más adecuadas para promover su reflexión sobre lo realizado. En segundo lugar, se señalan dos limitaciones principales para el desarrollo de la función formativa de la evaluación: la falta de referentes precisos para valorar las tareas y las formas en que se realiza la retroalimentación a los estudiantes.

En el año 2014, un equipo de investigadores del Instituto Nacional de Evaluación Educativa (INEEd) de Uruguay desarrolló otro estudio de casos sobre prácticas de evaluación en las aulas, en 6° de Primaria y 1° de Educación Media Básica. En su Informe Ejecutivo (INEEd, 2015) se señala la existencia de diferencias entre primaria y media con relación al uso de la evaluación con fines formativos. El estudio concluye que, mientras los maestros de primaria que participaron del estudio se refieren a la evaluación como un proceso cuyo propósito principal sería ayudar a los estudiantes a avanzar en los logros educativos esperados, los profesores de educación media se enfocan más en el señalamiento de errores y la calificación. El estudio señala expresamente que:

> *"[En primaria] en algunos casos predominaría la regulación retroactiva o remedial, de acuerdo a la cual la evaluación tendría un rol formativo en el proceso de aprendizaje de los niños: volver sobre los errores antes de seguir avanzando. A su vez, varios maestros entrevistados afirman que la evaluación les aporta elementos valiosos para revisar sus propias decisiones de enseñanza. Muy pocos vinculan de manera directa la evaluación con la tarea de calificar..."*
>
> *"[En secundaria] ...la función pedagógica de la evaluación, vinculada a la regulación de los aprendizajes, ocupa en la mayoría de los testimonios de docentes de educación media un lugar subsidiario. El papel de la evaluación en la tarea de enseñar tampoco tiene una presencia marcada en el discurso de los profesores. El error es visualizado en algunos casos como una "falla" que debe ser "corregida", siendo escasos los testimonios que dan cuenta de regulaciones que promuevan la revisión de lo propuesto por el docente o lo producido por los estudiantes como parte inherente al proceso mismo de enseñar y aprender"* (INEEd 2015:23).

Si bien en este estudio se concluye que hay ciertas diferencias entre las formas de trabajo prevalecientes en primaria y media, se pudo observar que las acciones formativas, cuando se producen, se centran fundamentalmente en los errores de los alumnos, tanto para subsanarlos como para incorporarlos en nuevas intervenciones.

Los hallazgos de los cuatro estudios de casos que acabamos de referir ponen de manifiesto la necesidad de revisar las concepciones y prácticas de evaluación en el aula en la región, con el fin de contribuir a que la misma cumpla más cabalmente con su función formativa de potenciar los procesos de aprendizaje de los estudiantes. Las evidencias recogidas en estos estudios muestran que los docentes manejamos distintas concepciones y preocupaciones al usar el término evaluación formativa y, sobre todo, que existen distintos énfasis con relación a la función que la misma cumple en el aprendizaje de los estudiantes.

La perspectiva que adoptamos en este libro, según explicaremos a lo largo del capítulo, es que la evaluación formativa es un proceso continuo, integrado naturalmente a las estrategias de enseñanza, de las cuales forma parte, y que tiene como propósito principal promover y hacer avanzar la reflexión, la comprensión y el aprendizaje de los estudiantes. Se concreta en el aula involucrando a cada estudiante, en la medida en que reflexiona sobre lo que realiza y aporta a la reflexión sobre las producciones de sus compañeros; y a cada docente, en tanto realiza devoluciones relevantes a sus estudiantes y revisa y ajusta sus propias estrategias y propuestas de enseñanza.

2. Evaluación formativa y para el aprendizaje: conceptos clave

Tal como señalamos en la introducción, en educación es tradicional la distinción entre evaluación formativa y evaluación sumativa. Esta distinción fue acuñada por Michael Scriven en 1967 para destacar el rol de la evaluación en el mejoramiento continuo del currículo, no de las pruebas en sí mismas (Wiliam, 2011). A las evaluaciones que tenían esta finalidad las llamó *formativas*, en tanto que denominó *sumativas* a las evaluaciones diseñadas para que las autoridades tomaran la decisión de si un currículo era apropiado como para ser adoptado en un sistema escolar. En 1969 Bloom aplicó esta distinción a las pruebas que se utilizan en las aulas, denominando como *formativas* a las evaluaciones cuya finalidad es dar devolución y orientación a docentes y estudiantes, como apoyo en cada etapa del proceso de aprendizaje (Wiliam, 2011). Las diferenció de las evaluaciones dirigidas a calificar a los estudiantes y, por tanto, a establecer a través de categorías o números, el logro alcanzado por cada estudiante en los contenidos de un curso. En 1971 Bloom y otros agregaron que un elemento central de la evaluación formativa es que sea empleada por el docente para tomar decisiones con respecto a su enseñanza, de modo de adaptarla en función de lo observado (Martínez Rizo, F., 2012).

En 1989, Sadler incorporó un elemento más: que la información recabada no solo debía ser utilizada por el docente, sino también por los propios alumnos, para mejorar su desempeño. El carácter formativo de la evaluación, por tanto, no es algo intrínseco a la actividad propuesta, sino que refiere al uso que se hace de la información recogida. Este uso involucra tanto al docente como a los propios estudiantes, a través de la autorregulación de sus procesos de aprendizaje (Sadler, 1989).

Más recientemente, un importante grupo de investigadores, integrado entre otros por Paul Black, Dylan Wiliam, Susan Brookhart y Rick Stiggins, incorporaron al concepto de evaluación formativa un componente de *"motivación del alumno para aprender"* (Martínez Rizo, F., 2012). Al respecto Brookhart (2008) señala que, si se aplica adecuadamente, la evaluación en el aula puede motivar a los estudiantes renuentes e involucrar a los desinteresados. Para ello es muy importante la capacidad que tengamos los docentes para *"ponernos en los zapatos de los estudiantes"*.

> *"El poder de la evaluación formativa reside en su enfoque de atención tanto a los factores cognitivos como a los motivacionales. Una buena evaluación formativa proporciona a los estudiantes información que necesitan para entender dónde están en su aprendizaje (el factor cognitivo) y desarrolla sentimientos de control de los estudiantes sobre su aprendizaje (el factor de motivación). Precisamente porque los sentimientos de autoeficacia de los estudiantes están involucrados, la retroalimentación aun bien intencionada puede ser muy destructiva si el estudiante lee el guión de una manera no intencional. Por ejemplo, a partir de una devolución, una estudiante puede pensar '¡yo sabía que era estúpida!'. No todos los comentarios bien intencionados tienen efectos positivos; la situación de comunicación tiene una gran importancia"* (Brookhart, 2008: 54).

En los últimos años algunos autores han reelaborado la distinción entre evaluación formativa y sumativa, utilizando las expresiones evaluación *del* aprendizaje y evaluación *para* el aprendizaje, indicando que en el primer caso se está intentando valorar lo que los estudiantes han logrado, en tanto que en el segundo la evaluación es parte intrínseca de los procesos de enseñar y aprender.

2.1. Dos lógicas diferentes

La evaluación formativa ha sido pues caracterizada de diversas maneras: por el tipo de instrumentos que utiliza, por el momento en que se realiza y por el foco en los procesos más que en los resultados. Sin embargo, lo

que define la distinción entre evaluación formativa y evaluación sumativa, y las constituye en dos formas de evaluación esencialmente distintas, es su finalidad. Según afirmamos más arriba, la evaluación formativa tiene como finalidad movilizar el aprendizaje y es parte de los procesos de enseñar y de aprender -como veremos un poco más adelante, es el puente entre ambos-. La evaluación sumativa tiene como finalidad certificar el aprendizaje, dar cuenta en forma pública de lo logrado por cada estudiante.

La distinción entre ambas, por tanto, no radica en los instrumentos, ni en el momento en que se realiza, ni en los aspectos evaluados. Un examen puede ser utilizado de manera formativa, si no va a ser calificado, para constatar lo logrado por los estudiantes hasta ese momento y decidir cómo avanzar. De la misma manera, la observación de los procesos de trabajo de un estudiante a lo largo del tiempo, puede ser utilizada como elemento de juicio para calificarlo o certificar su nivel de logro.

Como consecuencia de esta diferencia en sus finalidades, ambas evaluaciones tienen lógicas diferentes y efectos diferentes en los estudiantes. Una misma actividad puede generar actitudes y comportamientos distintos en los estudiantes, según si será calificada o no. Shepard (2006) señala que el uso excesivo de las calificaciones, que se incrementa especialmente al pasar de primaria a secundaria, tiene como efecto principal que las notas y no el aprendizaje, se conviertan en el objetivo central de los estudiantes.

> *"La evaluación sumativa y la calificación constituyen una seria amenaza para los objetivos de aprendizaje declarados por la evaluación formativa. De acuerdo con descubrimientos de la literatura motivacional y de estudios de maestros y estudiantes, las prácticas de las calificaciones pueden minar el proceso de aprendizaje de varias maneras. En primer lugar, las pruebas y las tareas calificadas comunican lo que es importante aprender. Si estas mediciones divergen de las metas del aprendizaje que se valora, entonces los estudiantes concentran su atención y esfuerzo solo en la porción calificada del currículo. Segundo, el uso de calificaciones como premio o como castigo puede socavar la motivación intrínseca de aprender. Tercero, a aquellos estudiantes para quienes los criterios de las calificaciones les parecen fuera de su alcance, éstas pueden reducir su esfuerzo y su ulterior aprendizaje"* (Shepard, L., 2006: 30).

Según vimos al caracterizar las prácticas de evaluación en la región, los docentes tenemos una clara tendencia a calificar continuamente las actividades de los estudiantes. La evidencia recogida muestra que solemos calificar todos los trabajos que proponemos a nuestros alumnos, así como otros aspectos de su desempeño. Esto seguramente responde a las múltiples deman-

das que recibimos, tanto de carácter institucional como social. Muchas veces la exigencia de las autoridades y de muchos padres respecto a que todos los trabajos escritos de los niños sean corregidos por el maestro, provoca que la misma se canalice a través de una nota, porque ello requiere menos tiempo del que implica hacer aportes y comentarios personalizados a cada estudiante. En la educación media el tiempo del docente es aún más limitado, porque la cantidad de estudiantes que cada uno tiene a su cargo es mucho mayor. De allí que la cantidad de trabajos que cada profesor debe revisar es mucho mayor que en primaria.

Más allá de las normas institucionales y del problema del tiempo, los propios alumnos nos reclaman que califiquemos los trabajos que nos entregan. Esto es parte de una cultura escolar, de la que formamos parte y que contribuimos a recrear con nuestras prácticas, porque no hemos sabido desarrollar otras maneras de informar a los estudiantes sobre sus desempeños. En cierto modo, nosotros mismos hacemos a los estudiantes "adictos" a las notas. Han aprendido a sacar conclusiones rápidas sobre su desempeño a partir de las mismas, más allá de que no les sean de mucha utilidad para saber qué hacer para avanzar en sus aprendizajes.

> *"El modelo de evaluación formativa y la investigación sobre la motivación sostienen que la calificación podría socavar la orientación de aprendizaje de los estudiantes. Por lo tanto, para conseguir que la evaluación formativa sea realmente para el aprendizaje, tal vez los maestros necesiten posponer otorgar calificaciones, o usarlas solo cuando el estudiante se autoevalúe, y como calificaciones hipotéticas que ayuden a los estudiantes a permanecer centrados en la retroalimentación sustantiva. Pero lo cierto es que los maestros deben evitar interrumpir y juzgar como si ya estuviera terminada la calidad del aprendizaje que aún está en marcha"* (Shepard, L., 2006: 37).

En la misma línea, Chappuis (2009) señala que asignar notas, puntajes o calificaciones a todo lo que el alumno hace, desvirtúa el sentido de las acciones realizadas con intención formativa, ya que los alumnos se sienten amenazados por las consecuencias que pueden tener en su trayectoria académica, por lo cual a menudo tratan de enmascarar sus equivocaciones, quitándonos ricos insumos para ayudarlos a avanzar.

Wiliam (2011) sostiene que la asignación y entrega de calificaciones a los estudiantes debería ser lo menos frecuente posible: una vez en cada bimestre en la educación media superior, una vez al año en la media básica, en forma completamente ocasional en la enseñanza primaria -de hecho, existen escuelas y propuestas pedagógicas que funcionan sin notas-. Kohn (1994, citado por Willian 2011) expresó el siguiente principio: *"Nunca se debe dar*

una calificación a los estudiantes mientras están en el proceso de aprendizaje", porque ni bien el estudiante recibe la calificación, el aprendizaje se detiene.

Por tanto, si queremos lograr un mayor involucramiento de los estudiantes con el aprendizaje, deberíamos buscar otros modos de informarles acerca de su desempeño, distintos de las calificaciones. Esto requiere desarrollar herramientas para la evaluación formativa y la realización de devoluciones adecuadas, tema sobre el que trabajaremos a lo largo de este capítulo. El problema de las calificaciones y la certificación lo analizaremos en el capítulo siguiente.

2.2. Valoración, orientación, retroalimentación y devolución

Para desarrollar una estrategia de evaluación formativa es importante partir de la distinción entre tres conceptos clave: valoración, orientación y retroalimentación o devolución (Wiggins, 1998). Los docentes tendemos a brindar a los estudiantes muchas *valoraciones*, pocas *orientaciones* y casi nada de *devoluciones* -en el sentido que vamos a explicar a continuación-. Las calificaciones son una forma de valoración. Expresan un juicio de valor que da a los estudiantes una pista acerca de qué tan cerca o lejos estuvieron de lo esperado. Pero como ni los criterios de logro ni el significado sustantivo de cada calificación suelen estar explícitos, el estudiante solo puede saber si está cerca o lejos de lo que pretendía el docente y si está mejor o peor que sus compañeros.

Muchas veces los docentes intentamos hacer evaluación formativa reemplazando las calificaciones por otras formas de valoración: felicitaciones y estímulos, así como distinto tipo de expresiones de satisfacción o insatisfacción con el trabajo (*'bueno', 'insuficiente', 'satisfactorio', 'no satisface aún')*. Especialmente en primaria, los docentes nos preocupamos por estimular a los estudiantes, haciéndoles comentarios positivos. Algunos ejemplos recogidos en nuestras entrevistas son los siguientes: "*Se les pone alguna observación. Por ejemplo al niño que ha trabajado muy bien se le pone una felicitación. Al que ha salido mal le pongo 'necesitas esforzarte un poco más', 'tu tarea no ha estado completa', 'muy bien, te felicito'*". Algunas maestras, incluso, consideran inconveniente hacerle notar al estudiante cuando ha hecho algo mal: *"Todo trabajo tiene validez. Si bien es cierto que hay niñas que no trabajan como uno espera, no es correcto decirles que su trabajo está mal.* [En esos casos] *yo pongo un cartelito 'Revisado' y luego cartelitos que dicen, 'Corrige ortografía', 'Entregue su trabajo ordenado', 'Dedíquese', 'Debe mejorar'"*. En la educación media, en cambio, los docentes suelen enfocarse más en indicar errores y asignar puntajes a las tareas.

Es importante darse cuenta de que las expresiones de valoración no le dan al estudiante información específica sobre qué y cómo mejorar su des-

empeño. Se limitan a hacerle notar qué tan cerca o lejos están de lo que el docente esperaba. Según Wiggins, las evaluaciones que hacemos los docentes se parecen al juego de "frío y caliente", en el que una persona elige un objeto de la habitación y la otra debe adivinar cuál es y dónde está, por ensayo y error, nombrando objetos. La primera persona le responde "frío", "tibio", "caliente", según qué tan lejos o cerca esté el objeto nombrado del objeto a descubrir. Así funciona la mayor parte de las valoraciones que hacemos habitualmente en el aula: como indicaciones ambiguas de qué tan cerca o lejos está el estudiante de lo esperado, pero sin ayudarle a comprender qué se espera ni por qué está cerca o lejos.

> "*...muchos educadores actúan como si no entendieran lo que es la devolución. Si tuviera que resumir lo que he visto en la última década en todo tipo de escuelas (públicas y privadas; de primaria, secundaria y universitarias; con y sin programas de exámenes estatales), debería reportar que muchos educadores parecen creer que la devolución significa brindar a los estudiantes mucha aprobación y un poco de desaprobación y consejo. En los salones de clase, la expresión más frecuente de la mal llamada devolución es: "¡buen trabajo!", "te felicito" o una frase equivalente. Por supuesto, es importante elogiar a los estudiantes porque los satisface y los anima. Los elogios te mantienen en el juego; pero solo la devolución real te ayuda a mejorar. La devolución te dice lo que hiciste y lo que no, y te permite realizar auto-ajustes. Por tanto, cuanto más auto-evidente sea la devolución, mayor será la autonomía que desarrollará el estudiante, y vice-versa*". (Wiggins, 1998: 46).

Asimismo, afirma Wiggins, los docentes tendemos a pasar directamente de la valoración a la orientación, es decir, de indicar insuficiencias o aspectos no logrados, a ofrecer pistas para mejorar el desempeño. El problema, afirma, es que muchas veces el estudiante no ha logrado comprender por qué su trabajo es insatisfactorio, por lo que la orientación no le sirve de mucho. "*El momento en que el estudiante* **comprende** *por qué una parte de su trabajo es inadecuada, es algo completamente diferente del momento en que el estudiante* **percibe** *que al docente no le gusta esa parte del trabajo*" (Wiggins 1998; 52, los subrayados son nuestros).

A partir de estas reflexiones, Wiggins propone como elemento central de la evaluación formativa el concepto de "devolución" o "retroalimentación". Wiggins utiliza el término anglosajón "*feedback*", que en castellano se traduce literalmente como **retroalimentación**, pero que para el ámbito de la educación preferimos traducir como **devolución**. La retroalimentación o devolución es información que el docente entrega a los estudiantes, por

diversas vías, que le ayudan a comprender el desempeño esperado y las diferencias con lo que efectivamente ha logrado. Dicho con otras palabras, la brecha entre lo aprendido y lo enseñado. Wiggins destaca especialmente que la devolución en lo posible no debería incluir valoraciones, sino más bien descripciones.

> *"La devolución es información sobre cómo una persona se desempeñó, a la luz de lo que intentó hacer –intento contra efecto, desempeño real contra desempeño ideal... La mejor devolución es altamente específica, directamente reveladora o altamente descriptiva de lo que realmente resultó, clara para el ejecutante, y disponible u ofrecida en términos de objetivos y estándares específicos"* (Wiggins, 1998: 46).

Recuadro 4.1: Un ejemplo paradigmático de devolución, el curso de soldadura de Ralph

"Consideren el siguiente ejemplo de un sistema que alienta y evalúa la auto-corrección por parte del estudiante. En su curso de soldadura, Ralph trabaja con 30 estudiantes a la vez y utiliza un sistema de devolución muy efectivo, que no requiere ninguna intervención del adulto. La primera tarea es muy concreta. Cada estudiante debe producir un ángulo de metal de 90 grados, siguiendo las especificaciones de la industria. Los estándares para esta soldadura están escritos en un papel, pero el sistema de devolución involucra algo más. Ralph les dice a los estudiantes que cuando crean que su soldadura cumple con los estándares deberán llevarla a una mesa, tomar un marcador y escribir su nombre en ella. Al hacer esto, los estudiantes estarán expresando, primero, que entienden los estándares, y segundo, que creen que su trabajo cumple con ellos. Sin embargo, hay una trampa. Sobre la mesa hay ángulos realizados en años anteriores, en un rango de excelentes a horribles. Cuando los estudiantes se acercan a la mesa pensando que han terminado la tarea y ven estos ángulos, piensan dos veces antes de escribir su nombre en el que ellos hicieron. Pude observar cómo algunos chicos miraban a su alrededor, antes de escurrirse hasta su estación de trabajo para rehacer su producto. Este es un sistema que, como todo buen sistema de devolución, funciona sobre la base de activar la capacidad que toda persona tiene para auto-evaluarse y auto-ajustarse cuando recibe una devolución clara y ve estándares claros" (Wiggins, 1998: 50).

Para ilustrar el concepto propone el ejemplo que se destaca en el Recuadro 4.1. El ejemplo es sumamente ilustrativo. Se trata de un dispositivo que permite a cada estudiante darse cuenta por sí mismo de la distancia entre lo que se esperaba que lograse y lo que efectivamente hizo. Y lo habilita a pensar en cómo mejorar su trabajo en forma inmediata. Obviamente, esto es relativamente fácil de hacer con producción de un objeto material, como es el caso de la soldadura de un ángulo de metal. Es más complejo con desempeños y producciones intelectuales, pero no imposible, como mostraremos en las páginas que siguen.

Tampoco debe derivarse del ejemplo la conclusión de que el docente no debe intervenir. Por el contrario, muchas veces la devolución la hace el propio docente, comentando el desempeño o el trabajo del estudiante. Lo importante es que estos comentarios no sean de tono valorativo, sino descriptivos y dirigidos a que el estudiante comprenda por sí mismo las insuficiencias de su trabajo y el modo de abordarlos. En estos casos la devolución se realiza a través del intercambio o de preguntas dirigidas a expandir la reflexión del estudiante durante el proceso de realización de las actividades que se le proponen, así como en el momento de revisar los productos parciales y finales de dichas actividades. Como veremos un poco más adelante, este tipo de devoluciones también puede provenir de los propios pares.

De acuerdo con Wiggins, una devolución adecuada tiene las siguientes características:

- Es frecuente y continua, se produce durante (en simultáneo con) el desarrollo de las actividades de evaluación, no solamente al final o después de las mismas.
- Se comunica principalmente mediante un lenguaje descriptivo.
- Brinda al estudiante evidencias concretas sobre su trabajo, que le permiten confirmar o no la pertinencia del mismo. El alumno puede comparar lo que hizo con lo que se esperaba.
- Ofrece ejemplos o modelos de tareas similares que satisfacen los criterios de calidad que debe alcanzar la misma, para que el propio estudiante saque conclusiones sobre el grado en que su desempeño ha alcanzado los aprendizajes esperados.
- Se describen diferentes niveles de logro en el desempeño de la tarea. Estas descripciones suelen presentarse en un dispositivo que se conoce como rúbrica o matriz de valoración.
- Permite que el estudiante se forme una imagen apropiada del desempeño que ha logrado, basada en criterios específicos. El estudiante podría incluso anticipar qué calificación hubiese recibido si la evaluación tuviera una finalidad certificativa.

- Incluye la posibilidad de que los estudiantes realicen autoevaluaciones y auto-ajustes.

Brookhart (2008), en tanto, alerta respecto de la necesidad de considerar la cantidad de información que el alumno puede procesar de manera efectiva. Ello implica que, debe estar apto para oírla y entenderla. Si el alumno no comprende el contenido y el contexto de la devolución que recibe, puede sentirse afectado negativamente desde el punto de vista emocional. Lo fundamental es centrarse en los procesos y procedimientos puestos en práctica para realizar la tarea. Esto le permite enfocarse en las estrategias de autorregulación o en sus habilidades como aprendiz. De esta manera la devolución promueve la actitud de que se puede mejorar en un nuevo intento.

3. La evaluación formativa como puente entre la enseñanza y el aprendizaje

Tal vez quien mejor ha expresado el rol de la evaluación formativa ha sido Wiliam, al definirla como el *puente* entre la enseñanza y el aprendizaje (2011: 46-50). Toda acción educativa es intencional: siempre que enseñamos los docentes nos proponemos lograr "algo" en los estudiantes. Este "algo" puede ser muy diverso, según vimos en el capítulo 2: memorizar datos, principios o hechos; aplicar fórmulas, algoritmos o procedimientos; explicar conceptos o propiedades; utilizar conocimientos para resolver situaciones y problemas nuevos; valorar situaciones para tomar decisiones basadas en información y valores; crear nuevos modelos, textos, productos o expresiones artísticas, entre otras. Lo importante a destacar es que nada de lo que enseñemos, ni el modo en que lo hagamos, garantiza que los estudiantes lo aprendan. Los docentes podemos enseñar e intentar que los estudiantes aprendan, pero no podemos asegurarlo. Dicho en otras palabras, los estudiantes no necesariamente aprenden lo que les enseñamos. Si fuese así no necesitaríamos evaluar ni calificar, alcanzaría con registrar en la libreta del docente lo que fue enseñado.

Los docentes no podemos predecir ni asegurar qué va a aprender cada estudiante, independientemente de cuán buena sea nuestra propuesta de enseñanza. Ni siquiera una excelente propuesta puede garantizar que todos los estudiantes aprendan. Asimismo, ante una misma propuesta de un mismo profesor, en la misma aula, distintos estudiantes comprenden y aprenden cosas distintas.

Esta cuestión está planteada de un modo muy ilustrativo en un reciente texto de la OCDE sobre el gobierno de los sistemas educativos (Burns, T.

y Köster, F., 2016: 21). Para explicar el problema de la complejidad en los sistemas en general, se distinguen tres conceptos: "simple", "complicado" y "complejo". **Simple** es algo que se puede hacer siguiendo una receta. Las recetas son fáciles de replicar. Para seguir una receta el conocimiento y la experiencia son importantes pero no indispensables. Las buenas recetas tienen casi siempre un buen resultado. **Complicado** es enviar un cohete a la luna. Las fórmulas y principios tienen un rol crítico. Tener éxito una vez aumenta la probabilidad de lograrlo en las siguientes. Se necesitan altos niveles de conocimiento especializado en diferentes disciplinas. Los cohetes son similares en varios aspectos y existe un importante grado de certeza de lograr el resultado esperado si las cosas se hacen adecuadamente. **Complejo** es educar a un niño. Las fórmulas, recetas y principios generales pueden ser útiles pero tienen una aplicación limitada. Educar a un niño le da a uno experiencia, pero no asegura el éxito con otro niño. El conocimiento especializado contribuye, pero no es indispensable ni suficiente para asegurar el logro de los propósitos. Cada niño es único y requiere de un abordaje individualizado. Es imposible eliminar un alto grado de incertidumbre sobre el resultado de la acción educativa.

Enseñar, entonces, es algo complejo. El resultado nunca está asegurado. Esto es algo que todos los docentes sabemos. Sin embargo, de esta constatación algunos docentes derivan la siguiente postura: debemos plantearnos objetivos de enseñanza, pero no de aprendizaje. Como no es posible garantizar que los estudiantes aprendan, la consecuencia lógica para muchos docentes es que el aprendizaje depende exclusivamente del esfuerzo de los estudiantes. Consideran que la labor del docente es "dar" la clase, pero no asegurar que los estudiantes aprendan. Desde esta postura se resisten a plantear metas de aprendizaje y, en términos generales, a las evaluaciones.

Desde nuestra perspectiva esa postura tiene una parte de verdad, pero no es toda la verdad. Es posible y necesario dar un paso más, desde centrarse exclusivamente en *"qué voy a dar"* y *"qué van a hacer los estudiantes"* hacia *"qué es lo que pretendo que los estudiantes aprendan y sean capaces de hacer"*. Aunque el aprendizaje no pueda ser garantizado, es indispensable hacer lo posible para lograrlo. Para ello la evaluación formativa tiene un rol crucial: justamente, su función principal es hacer visibles las brechas entre lo que el docente enseña y lo que los estudiantes aprenden. En otras palabras, un buen docente, además de enseñar bien, debe ser un buen evaluador. La evaluación es el puente entre enseñanza y aprendizaje, que evita que sean procesos paralelos que nunca se encuentran. Una capacidad clave de un buen docente es la de percibir y valorar la distancia entre lo que se propuso lograr -sus intenciones educativas- y lo que realmente alcanzó cada estudiante -aprendizaje-. A partir

de allí puede tomar acciones de dos tipos. Por un lado, modificar y ajustar su propuesta de enseñanza. Por otro lado, hacer devoluciones a los estudiantes que les permitan comprender la brecha existente entre lo enseñado y lo que han aprendido. En este proceso, como veremos enseguida, el estudiante debe asumir su propio rol como responsable de su aprendizaje.

4. Las cinco estrategias centrales de la evaluación formativa

En la evaluación formativa existen tres actores clave y tres procesos clave (Wiliam 2011: 45-46). Los actores clave son el docente, el estudiante y el grupo de pares[4]. Los procesos clave son: i) clarificar y compartir las intenciones educativas de un modo que los estudiantes puedan comprenderlas -adónde deberían llegar-; ii) generar evidencias acerca de qué están aprendiendo -dónde están los estudiantes con relación a las intenciones educativas-; iii) realizar devoluciones y ofrecer orientación que permita a cada estudiante ajustar su desempeño y continuar aprendiendo -cómo avanzar hacia el "lugar" deseado-[5].

Figura 4.2: Cinco estrategias de la evaluación formativa

	Hacia dónde va el estudiante	Dónde está el estudiante justo ahora	Cómo avanzar hacia dónde queremos ir
Docente	Clarificar y compartir las intenciones educativas y criterios de logro	Diseñar y llevar adelante actividades que ofrezcan evidencia de qué está aprendiendo cada estudiante	Proporcionar devoluciones que movilicen el aprendizaje en la dirección deseada
Pares	Comprender y compartir las intenciones educativas y los criterios de logro	Activar a los estudiantes como fuente de aprendizaje para sus pares	
Estudiante	Comprender las intenciones educativas y los criterios de logro	Activar a cada estudiante como responsable de su propio aprendizaje	

Fuente: Tomado y adaptado de Wiliam (2011: 46)

4. Desde la perspectiva de la evaluación en situaciones auténticas, podríamos agregar un cuarto actor, los destinatarios externos de la producción de los estudiantes, cuando los hay.
5. Como se puede apreciar, Wiliam se enfoca en los estudiantes y no menciona la revisión de la propuesta de enseñanza por parte del docente.

A partir de la combinación de estos actores y procesos, Wiliam define lo que considera las cinco estrategias clave de la evaluación formativa como puente entre la enseñanza y el aprendizaje, que son las siguientes:

1. Compartir, clarificar y comprender las intenciones educativas y criterios de logro
2. Diseñar y llevar adelante actividades y tareas que ofrezcan evidencia de lo que cada estudiante está aprendiendo
3. Proporcionar devoluciones que movilicen el aprendizaje en la dirección deseada (que hagan avanzar el aprendizaje)
4. Activar a los estudiantes como fuente de aprendizaje para sus pares
5. Activar a cada estudiante como responsable de su propio aprendizaje

En lo que sigue de este capítulo nos enfocaremos en ofrecer ejemplos y formas de trabajo concretas para poner en práctica cada una de estas estrategias en el aula.

4.1. Clarificar, compartir y comprender las intenciones educativas y criterios de logro

Una de las principales dificultades que tienen muchos estudiantes para aprender es que no comprenden suficientemente qué es lo que el docente espera de ellos. En algunas ocasiones, incluso, hay estudiantes a los que les va mal en una prueba porque piensan que les están pidiendo algo más complejo de lo que en realidad se pretende. Como fue señalado antes, los docentes tendemos a enfocarnos más en lo que vamos a "dar" -los temas- que en lo que los estudiantes van a aprender. Necesitamos preocuparnos más por clarificar, discutir y ayudar a los estudiantes a comprender lo que esperamos que logren en cada curso o unidad.

Lograr que los estudiantes comprendan qué es lo que se espera de ellos y cuáles son las metas de aprendizaje, es algo difícil de llevar a la práctica, por la sencilla razón de que no alcanza con establecer una meta u objetivo para cada clase o unidad. Muchas veces se intenta explicitar las metas de aprendizaje de un modo formal, escribiendo "objetivos de aprendizaje". La mayor parte de las veces esto termina siendo un simple ritual: ni los estudiantes comprenden los objetivos escritos, ni el docente los utiliza realmente en sus clases. En palabras de Wiliam (2011: 56), *"el profesor escribe el objetivo en el pizarrón, los estudiantes copian el objetivo en su cuaderno, y luego el objetivo es ignorado durante el resto de la clase"*.

La cuestión central es cómo hacer para que las intenciones educativas y los aprendizajes esperados, sean comprendidos y significativos para los estudiantes. No se trata de una receta a cumplir (poner siempre un objetivo de

aprendizaje), sino de un proceso de explicitación, discusión y comprensión. Muchas veces la calidad de un trabajo o desempeño es difícil de definir con palabras y tiene un componente fuertemente intuitivo (Pirsig, 2004). No puede expresarse del todo en forma abstracta. Por este motivo, afirma Wiliam, no alcanza con definir objetivos. Es necesario, además y sobretodo, desarrollar en los estudiantes el "olfato" para identificar buenas producciones. Esto constituye para el docente una actividad artesanal e interactiva, más que de escritura formal de objetivos.

A continuación ofrecemos tres pistas concretas para trabajar en el aula, orientadas a ayudar a los estudiantes a comprender las características de un buen trabajo, comprender lo que se espera de ellos y revisar en forma reflexiva sus propias producciones:

a. Elegir **cinco trabajos** realizados por estudiantes (ej. informes de laboratorio, narraciones, ensayos de historia o filosofía), **de distinta calidad** (alguno muy bueno, alguno muy malo y algunos intermedios) y pedirles que, en equipos, los lean y ordenen según su "calidad" de 1 a 5 (sin anticiparles la valoración del docente). Una vez ordenados los trabajos, cada equipo debe discutir las razones por las que unos son mejores que los otros y ponerlas por escrito. El paso siguiente es compartir en el grupo completo el modo en que cada equipo ordenó los trabajos y las razones para ello, dando lugar a un intercambio colectivo. Como resultado de este proceso se puede construir una rúbrica (sobre lo cual abundaremos más adelante en este capítulo). Esta dinámica de trabajo puede ser aplicada a una amplia gama de desempeños y productos. Por ejemplo, en una clase de inglés se puede aplicar a la escucha de grabaciones de estudiantes leyendo en voz alta para discutir en cuáles es mejor la pronunciación y por qué, o a cuentos escritos por los propios estudiantes. Como se puede apreciar, esta es una forma de construir dispositivos de devolución semejantes al ejemplo del curso de soldadura propuesto por Wiggins (Recuadro 4.1), pero aplicados a producciones intelectuales.
b. Pedir a cada estudiante que, individualmente o en equipo, compare su propio trabajo con dos o tres de los mejor logrados, seleccionados por el docente. Como resultado de esta actividad cada estudiante debe **identificar las principales debilidades e insuficiencias de su trabajo y hacer una segunda versión del mismo.**
c. Pedir a los estudiantes que **elaboren preguntas o problemas para una prueba** sobre un tema que haya sido trabajado en clase. Solicitar a los estudiantes que elaboren preguntas le permite al docente

darse cuenta de qué creen los estudiantes que han estado aprendiendo (Wiliam 2011: 68). Al mismo tiempo, ayuda a los estudiantes a expresar y confrontar qué tipos de desempeños visualizan como las intenciones educativas del docente. Una vez elaboradas las preguntas, se hace una discusión colectiva acerca de qué preguntas son buenas y cuáles no, y por qué.

Más allá de estas pistas concretas, dos consideraciones adicionales son importantes con respecto a la clarificación de las intenciones educativas para los estudiantes. En primer lugar, que no siempre es necesario ni conveniente explicitar las metas u objetivos **antes** de trabajar un tema o de iniciar una unidad o clase. Muchas veces la motivación y el efecto formativo de una actividad requiere que los estudiantes no sepan hacia dónde van o qué pretende el profesor (obviamente, nos referimos al tiempo de la enseñanza y a la evaluación formativa, no a una situación de prueba). De todas formas, aún en estos casos, es importante que al final del proceso quede claro cuál fue el sentido de la actividad realizada, por qué se hizo y qué se esperaba como aprendizaje (Wiliam, 2011: 57).

En segundo lugar, una tensión que los docentes debemos tener en cuenta y manejar según nos parezca más apropiado, es la relativa al tipo de lenguaje, formal o amigable, que usamos para expresar las logros de aprendizaje esperados. A lo largo del curso y en distintos momentos deberíamos utilizar ambos tipos. Es importante, por un lado, facilitar que los estudiantes entiendan lo que se espera que aprendan, empleado un lenguaje sencillo, directo y amigable. Pero, simultáneamente, es importante que aprendan a comprender y utilizar el lenguaje formal y abstracto propio de la disciplina (Wiliam, 2011: 65).

4.2. Diseñar y llevar adelante actividades que ofrezcan evidencia de lo que cada estudiante está aprendiendo

En el apartado anterior analizamos la necesidad de llevar adelante acciones dirigidas a que los estudiantes comprendan hacia dónde van, es decir, qué pretendemos que logren a través de lo que les estamos enseñando. El segundo desafío clave que tenemos como docentes es darnos cuenta, con la mayor frecuencia posible, dónde están nuestros estudiantes en cada momento. En este apartado nos proponemos aportar algunas ideas y pistas de trabajo para ello.

El desafío es complejo. Como solemos estar a cargo de muchos estudiantes, es difícil darse cuenta de qué es lo que está entendiendo cada uno. En el discurso pedagógico contemporáneo se pone mucho énfasis en

la importancia de adaptar la enseñanza a las necesidades y estilos de cada estudiante, y de abrir espacio para la diversidad de inteligencias y formas de aproximarse a los saberes. Hacerlo es bastante más complejo que decirlo. Muchas veces los docentes renunciamos a intentarlo, damos nuestra clase y hacemos de cuenta que todos los estudiantes nos siguen y entienden lo que les explicamos.

El problema es que solemos planificar lo que vamos a exponer y explicar, así como las actividades que van a realizar los estudiantes, pero por lo general no planificamos demasiado las instancias y preguntas que nos permitirán verificar quiénes están comprendiendo y qué es lo que están comprendiendo. A lo sumo formulamos las típicas preguntas del estilo '*¿entendieron?*', '*¿quedó claro?*' o '*¿alguien no entendió?*', ante las cuales es excepcional que algún estudiante responda que no entendió. Cuando lo hace, tenemos un problema, porque debemos detener el desarrollo de la clase y perdemos la atención de la mayor parte del grupo. Reiteramos la explicación con alguna modificación, pero generalmente el estudiante sigue sin entender. Luego de un par de intentos, ambos, docente y estudiante, nos damos por vencidos.

Un componente central de la evaluación formativa es idear preguntas y tareas que nos permitan explorar de modo más o menos continuo qué es lo que los estudiantes están comprendiendo. En este apartado nos vamos a enfocar en lo que el docente puede hacer en el aula, es decir, durante la enseñanza (en el capítulo 2 ya hemos analizado cómo elaborar actividades escritas de evaluación que nos aporten evidencia sobre los procesos cognitivos que los estudiantes están desarrollando). Más allá de las instancias "formales" de evaluación (pruebas y similares), que pueden tener finalidad formativa o certificativa, la evaluación formativa, es decir, la constatación de la convergencia y/o de las brechas entre enseñanza y aprendizaje, es parte esencial de la labor docente en el aula.

El primer paso que necesitamos dar es revisar cómo suelen desarrollarse nuestras clases. ¿Qué hacemos generalmente en el aula? Diversos estudios, entre ellos uno que tuvimos oportunidad de realizar entre 2009 y 2011, filmando clases de 6º de primaria en cinco países de la región[6], así como

---

6. El estudio referido fue realizado durante los años 2011 y 2012 en forma conjunta por el Instituto de Evaluación Educativa de la Universidad Católica del Uruguay (IEE-UCU) y la Organización de Estado Iberoamericanos (OEI, Oficina de Buenos Aires). La denominación del estudio fue **Prácticas de Enseñanza en América Latina – Banco de Registros Digitales**. Se filmaron 142 clases de 6º de Primaria en las ciudades de Asunción (Paraguay), México DF (México), Cali (Colombia), Santiago de Chile (Chile) y Córdoba (Argentina). A partir de estas filmaciones se generó un banco de registro digitales y, posteriormente, se produjeron 250 videos cortos (de

nuestra experiencia de observación de aulas, nos indican que el formato predominante en el desarrollo de las clases es como sigue. El docente conduce la explicación en torno a un tema; a lo largo de su presentación hace preguntas que algunos estudiantes responden; el profesor "recoge" algunas respuestas, en general aquellas que más se adecuan a la exposición que viene desarrollando (a veces anota algunas palabras en el pizarrón); apoyándose en dichas respuestas, continúa adelante con su explicación. Las preguntas suelen ser de respuesta muy breve (se responden con una palabra o una frase). Muchas veces se parecen a "adivinanzas" que los estudiantes responden todos al mismo tiempo, en una suerte de coro. En algún momento el docente hace un resumen que los estudiantes deben escribir en sus cuadernos.

Una variante de este formato, que se observa principalmente en la educación media, consiste en que el docente "interroga" con más detalle a algunos estudiantes, que se supone estudiaron antes el tema en su casa. El involucramiento de los estudiantes en la temática en curso en la clase es parcial. No todos los estudiantes se involucran, solo algunos. Muchos lo hacen en forma intermitente, se conectan y desconectan de la clase durante el desarrollo de la misma.

En cualquiera de las dos variantes, los docentes tendemos a seguir nuestro "guión" explicativo definido previamente, con la mayor fidelidad posible, eventualmente con algunas ramificaciones en función de las intervenciones de los estudiantes. Pero, por lo general, no se producen procesos propiamente reflexivos. No suele haber discusiones de cierta profundidad. La mayoría de las preguntas requieren recordar algo muy breve. Muy pocas preguntas están dirigidas a que los estudiantes realicen inferencias, valoraciones, reflexiones o tomas de posición.

Wiliam (2011: 77) menciona un estudio realizado en los Estados Unidos, en el que fueron registradas mil preguntas realizadas por docentes en el aula. De ellas, el 57% eran preguntas de organización, tales como "¿Quiénes terminaron la tarea?" o "¿Trajeron sus libros?". Un tercio de las preguntas requerían únicamente recordar hechos o datos como, por ejemplo, "¿Cuántas patas tiene un insecto?". Solo el 8% de las preguntas registradas requerían que los estudiantes reflexionaran, realizaran inferencias o generalizaciones como, por ejemplo, "¿Por qué un pájaro no es un insecto?".

---

4 minutos de duración), con el fin de generar un sitio web. Lamentablemente esta última etapa nunca pudo ser concretada.

a. Dos clases distintas

Con el fin de ilustrar las dinámicas de clase, en los Recuadros 4.2 y 4.3 incluimos dos transcripciones de clases filmadas en video en 6° año de primaria en el estudio antes mencionado. El registro incluido en el Recuadro 4.2 contiene un diálogo continuo entre maestra y estudiantes. La maestra pregunta sobre qué características deben tener los avisos y cómo se presentan. Distingue entre avisos informales y formales. Utiliza como ejemplo una comunicación de la Dirección de la escuela a los padres. Trabaja "dialogando" directamente con todo el grupo, pero no promueve el diálogo <u>entre</u> los estudiantes. Las intervenciones de los estudiantes son utilizadas para reafirmar los conceptos que la docente quiere presentar: aviso, destinatario, mensaje, remitente.

Recuadro 4.2: Análisis de un aviso (6º grado de primaria)

A partir de un intercambio sobre "cómo debe ser un avisito", analizan una comunicación de la Dirección de la escuela a los padres. La leen en voz alta, primero una niña y luego todos a coro, mientras la docente va anotando al costado de cada una de las partes del texto, las palabras fecha, destinatario, mensaje y remitente.

M: "¿Cómo debe ser un avisito?"
N: "Informativo"
N: "Sencillo"
N: "Claro"
N: "Breve"
M: "¿Cómo se presentan esos avisos?"
N: "A través de la televisión"
N: "Internet"
N: "Volantes"
N: "Prensa escrita"
N: "Pasacalle"
M: "Esos avisos que presentamos así, que vemos en la calle, como pasacalles, ¿qué viene a ser? Un aviso informativo, que da noticia momentánea [la maestra responde su propia pregunta]. Hay también avisos en donde es más formal".

Texto en el pizarrón:
6 de octubre de 2009.
Señores padres:
La dirección de la institución, comunica que el 9 octubre del corriente año a las 10 horas, se realizará una reunión general del año escolar en el local de la escuela.
Atte,
La Dirección.

M: "Vamos a ver sus partes. Primeramente tenemos la fecha ¿Luego qué dice?"
N: "Señores padres"
M: "Viene a ser destinatario. Luego dice.."
Niños leen a coro el cuerpo del texto.
M: "¿Qué viene a ser eso?"
N: "El mensaje"
N: "Fecha, destinatario, mensaje, remitente".

Fuente: Estudio de Prácticas de Enseñanza en América Latina: Banco de registros digitales, 2012. Caso 17.

Como resulta evidente, las preguntas de la maestra no están dirigidas a promover la reflexión ni a indagar qué están comprendiendo los estudiantes, sino a servir de apoyo al itinerario temático que ha planificado para la clase. Se trata de preguntas cerradas, cuya respuesta está predefinida y consiste en una sola palabra. Las preguntas no habilitan la posibilidad de repreguntar o promover que los participantes busquen alternativas de respuesta. La docente no genera una discusión ni una instancia de participación reflexiva, sino que utiliza las respuestas de los estudiantes para reforzar el guión expositivo que tiene previamente definido. Tampoco promueve la interacción entre los estudiantes. En contraposición al anterior, el Recuadro 4.3 muestra una clase en la que la docente desarrolla un proceso de diálogo reflexivo con los distintos equipos de trabajo.

Recuadro 4.3: Resolución de problemas de razonamiento lógico (6º grado de primaria)

M: "Bueno chicos, hoy vamos a trabajar unos problemitas que les traje de un libro muy lindo de acertijos y adivinanzas y nos hemos agrupado como para trabajar en grupos, ¿de cuántos? De parejas. Antes de resolver el problema, ¿qué tenemos que hacer?"
N: (a coro): "Leerlo atentamente"
M: "Leerlo atentamente. ¿Después?"
N: "Determinar qué datos son importantes".
M: "Muy bien. Qué datos son importantes para resolver ese problema. ¿Y después qué hacemos?"
N: "Resolver"
M: "¿Qué podemos hacer para resolver? ¿Qué utilizamos?"

Recuadro 4.3: Resolución de problemas de razonamiento lógico (6º grado de primaria)

N: "Gráficos, operaciones, o dibujos".

....

Presenta en el pizarrón

El árbol genealógico

En la familia Del Pino las relaciones de parentesco son un poco complicadas:

- Román y Miguel son hijos de Lucas.
- José solo tiene dos hijos.
- Lucas no es el padre de Claudio.
- Carlos es el hermano de Lucas.
- José es el abuelo de Claudio.
- Lucas es hijo de José.

¿Quién es el padre de Claudio?

Una niña lee el problema presentado.
Los alumnos trabajan en parejas.

Intervenciones de la maestra en los diversos equipos:
M: "¿Vos para resolver un problema que tenés que buscar?"
N: "Quién es el padre de Claudio"
M: "Esa es la pregunta que te hace el problema. ¿Qué te da para que vos encuentres esa respuesta? Que es muy importante para resolver el problema"
N: Duda y la maestra lo ayuda.
M: "Los datos. Sin los datos no podemos resolver el problema"
Se dirige a todo el grupo:
M: "A partir de los datos van a encontrar la respuesta. ¿Qué pueden hacer para verlo, visualizarlo a eso? ¿Qué pueden utilizar?"
N: "Un dibujo"
M: "Bueno; un dibujo, un gráfico. Prueben con eso"

....

M: "¿Dónde ubicarías en este gráfico a Carlos? Si Lucas es el hermano de Claudio, pensá"

....

N: "Seño, ya lo sacamos: Carlos es el padre de Claudio"
M: "¿Lo saben justificar?"

Recuadro 4.3: Resolución de problemas de razonamiento lógico (6º grado de primaria)

N: "Sí"

M: "Perfecto"

.... Luego de un tiempo de trabajo en parejas, la maestra vuelve a trabajar con todo el grupo.

M: "¿Quién se anima a pasar al frente y demostrarme y justificar la respuesta en el pizarrón?"

Pasa una alumna al frente:

N: "Nosotros lo que hicimos fue ir viendo esto"

M: "¿Qué eran esos?"

N: "Las consignas, las pistas"

M: "Los datos más importantes que me da el problema".

N: "Y fuimos fijándonos la familia de cada uno y pusimos que José [escribe en el pizarrón] tenía 2 hijos. Buscamos los hijos de José y nos dio que eran Lucas y Carlos. Entonces los representamos en un gráfico [escribe en el pizarrón lo que va diciendo]. Lucas tiene 2 hijos que son Román y Miguel. Y de ahí buscamos y nos dio que Carlos era el padre de Claudio; Claudio era el hijo de Carlos".

M: "Muy bien".

M: "¿Alguien lo hizo de otra forma?"

N: "Nosotros"

M: "A ver pasa Brian, así nos cuentan cómo lo hicieron"

N: "Nosotros buscamos en los datos quién era el padre de Claudio y nos dio que era Carlos, por la generación de la familia".

M: "¿A ver cómo lo hicieron? ¿Hicieron un gráfico?"

N: "No, hicimos la familia" [Dibuja en el pizarrón y explica al mismo tiempo. Representa cada miembro de la familia con un dibujo]. Este era José, era el abuelo de Claudio. Acá decía que tenía 2 hijos: eran Lucas y Carlos. Buscamos en el cuaderno y nos dio que Lucas no era el padre de Claudio"

M: "¿Entonces qué hicieron, descartaron a Lucas como padre?"

N: "No, no era. Pero sí que tenía dos hijos que eran Román y Miguel. Y acá nos dio que Carlos era el padre de Claudio. Entonces por la generación Román y Miguel son primos de Claudio. Y Lucas y Carlos, hermanos"

José / Lucas / Carlos / Román / Miguel

M: "Muy bien; o sea que ellos fueron viendo los niveles de parentesco que tenían"

M: ¿"Ellos lo resolvieron igual?"

N: "No" [a coro]

Recuadro 4.3: Resolución de problemas de razonamiento lógico (6º grado de primaria)

M: "Ven que cada uno puede utilizar una estrategia, un modo de resolver el problema como mejor lo crea conveniente"

......................[Pasa otra pareja]

N: "Elegimos, haciendo el árbol genealógico" [dibuja en el pizarrón y va explicando al mismo tiempo]. "Hicimos el árbol, pusimos a Claudio en el medio y luego fuimos probando uno por uno. Pusimos a Lucas, pero decía que él no podía ser porque tenía 2 hijos. Y acá pusimos a Miguel y a Román. Y acá Carlos, porque acá decía que el padre era José, entonces el abuelo de Claudio era José"

M: "Perfecto. ¿Alguien más lo hizo de alguna otra forma?

....

M: "¿Cuál fue el dato clave para ustedes que les hizo decidir que sí o sí Carlos era el papá de Claudio?"

N: "De que Lucas no era el papá de Claudio"

N: "Y José era el abuelo"

N: "Ese, y además que quedaba solamente Carlos sin hijo y ahí queda Claudio solo"

....

M: "¿Hay algún dato que ustedes podrían haber agregado? Mirando allá los esquemas, ¿qué podrían haber agregado? Algún otro dato más..."

N: "Sobre los primos"

M: "Muy bien, ese es un dato ¿qué podrían haber escrito como dato acá?"

N: "Román y Miguel son primos de Claudio"

M: "Muy bien, Román y Miguel son primos de Claudio. Entonces les hubiera dado un dato más para llegar más fácil a la respuesta.

M: "Y por ejemplo, si yo les hubiera sacado este dato [señala en el papelógrafo: Carlos es el hermano de Lucas], ¿hubieran llegado fácilmente a la respuesta?"

N: "No" [a coro]

M: "¿Por qué?"

N: "Porque ahí dice que es el hermano de Carlos"

M: "¿Y en algún otro lugar aparece el nombre de Carlos?"

N: "No" [a coro]

M: "Bueno, muy bien. Han trabajado hermoso".

Fuente: Estudio de Prácticas de Enseñanza en América Latina: banco de registros digitales, 2012. Caso 237.

En la primera parte del diálogo establecido por la maestra las preguntas son cerradas y de evocación, al igual que en el caso anterior. Pero en esta clase es claro que la finalidad de las preguntas es activar la evocación de los conceptos necesarios para enfrentar organizadamente la situación propuesta por la docente. A medida que avanza la clase, se observa que la docente genera instancias de intercambio, ya que organiza la resolución de la actividad en parejas. Asimismo, interviene con preguntas focalizadas en las tareas que va desarrollando cada dupla, de manera ajustada al proceso que va haciendo cada una. Las preguntas son abiertas y están dirigidas a ayudar a cada pareja a avanzar en la tarea, según dónde se encuentran. A través de estos aportes, los alumnos van teniendo pistas sobre cómo continuar e incluso, la posibilidad de preguntar cosas nuevas. De esta manera van tomando decisiones durante el proceso de resolución, disponiendo de la oportunidad de equivocarse y aprender a partir de sus errores y también de los de sus compañeros.

Por último, la maestra solicita que cada dupla explique el procedimiento seguido para resolver la situación planteada, mostrando la existencia de diferentes caminos para hacerlo. Esto hace que cada pareja en cierto modo "abra" -explicite- sus procesos de pensamiento y que todo el grupo reflexione en forma conjunta. En esta instancia es interesante observar cómo la docente promueve la reflexión de los estudiantes, planteando posibles cambios que se podrían haber hecho en la propuesta inicial.

b. Las preguntas del docente como herramienta para promover y evidenciar el aprendizaje

Las preguntas que hace el docente tienen una importancia crucial para definir el tipo de aprendizaje que se promueve en las aulas. Según Wiliam (2011: 79), solo hay dos buenas razones para que un docente haga preguntas en clase: a) para promover la discusión y la reflexión de los estudiantes; b) para obtener información que le permita saber qué están entendiendo y cómo seguir adelante. En otras palabras, carecen de sentido tanto las preguntas retóricas que están dirigidas a apoyar el desarrollo del discurso del docente, como las preguntas orientadas a "tomar la lección", en las que el estudiante debe reproducir conceptos, principios o definiciones (salvo que la intención educativa del docente esté orientada a la memorización).

La cuestión de las preguntas y actividades escritas orientadas a promover la reflexión de los estudiantes y a generar evidencia sobre sus procesos cognitivos, fue trabajada en los capítulos 2 y 3, al analizar los tipos de consignas y situaciones en las evaluaciones escritas. En este apartado nos detendremos, en el contexto de los casos que acabamos de ver, en el análisis de las

preguntas que hacemos oralmente en el aula durante la enseñanza, en tanto herramientas de evaluación formativa.

Dada la brecha inevitable entre lo que enseñamos y lo que los estudiantes aprenden, es necesario realizar constantemente preguntas que nos permitan explorar qué están entendiendo los estudiantes. Todos entienden "algo", aunque sea errado, y todos entienden de un modo parcial, cada estudiante entiende aspectos distintos. Sin embargo, en general hay "grupos" de comprensiones comunes dentro de la clase, no es que haya tantas comprensiones como estudiantes.

En este marco, las preguntas del docente juegan un doble rol: generan reflexión en los alumnos y le abren al docente "ventanas" hacia el pensamiento de los estudiantes. Las buenas preguntas son clave para la enseñanza pero, a la vez, algo difícil de lograr. Tres pistas de trabajo para formular preguntas que provoquen reflexión y comprensión, son las siguientes (Wiliam 2011: 86 y ss.):

a. Evitar las preguntas que se responden con 'sí' o 'no', o con una sola palabra. Reemplazarlas por preguntas del tipo 'por qué'. Algunos ejemplos propuestos por el autor son los siguientes: en vez de preguntar '*¿un pájaro es un insecto?*", deberíamos preguntar '*¿por qué un pájaro no es un insecto?*'; en lugar de preguntar '*¿un cuadrado es un trapecio?*'; deberíamos preguntar '*¿por qué un cuadrado es un trapecio?*'; en lugar de preguntar si '*pensar*' es un verbo regular, deberíamos preguntar '*¿por qué "pensar" es un verbo irregular?*'; en vez de preguntar si la fotosíntesis es una reacción endotérmica, deberíamos preguntar '*¿por qué la fotosíntesis es una reacción endotérmica?*'; y así sucesivamente.
b. Cambiar las preguntas que piden a los estudiantes que den una definición por preguntas que les requieran establecer una comparación o dar un ejemplo. Por ejemplo, en vez de preguntar '*¿cómo era la vida durante el apartheid?*', deberíamos preguntar '*¿en qué se diferenciaban las vidas de blancos y negros durante el apartheid?*'; en lugar de preguntar '*¿qué es un mamífero?*', podríamos preguntar '*¿por qué un murciélago es un mamífero y un pingüino no lo es?*'.
c. Utilizar, en lugar de preguntas, afirmaciones controversiales ante las cuales los estudiantes deban tomar posición y argumentarla. Por ejemplo, en vez de preguntar qué país fue responsable del inicio de la Primera Guerra Mundial, se podría plantear la afirmación '*Rusia fue la principal responsable del inicio de la Primera Guerra Mundial*' y ponerla a discusión entre los estudiantes. Lo mismo con la afirmación '*todos los cuadrados son rectángulos*'.

También es importante la forma de plantear las preguntas en el aula. Algunos docentes tenemos tendencia a dirigir nuestras preguntas a un grupo reducido de alumnos, que habitualmente nos "siguen". El resto suele "desconectarse" de la clase: algunos atienden un rato, levantan la mano para responder algo que saben, y luego se dedican a otra cosa porque saben que ya nos les volveremos a preguntar a ellos. Otras veces, cuando no obtenemos la respuesta que deseamos, solemos recurrir a algún estudiante que sabemos nos dará una respuesta apropiada, lo cual nos permitirá seguir adelante con nuestro guión. Un buen docente, afirma Wiliam, se preocupa por mantener a todos los estudiantes involucrados en la clase. Una parte importante de las desigualdades en el aprendizaje obedece, sencillamente, a que algunos se involucran y participan en las clases, y otros no. Para afrontar esta situación, es necesario pasar de una forma de preguntar centrada en obtener la respuesta correcta a una forma de preguntar centrada en comprender la forma de pensar del estudiante. Para ello pueden ser de ayuda algunas prácticas concretas:

a. elegir en forma aleatoria a los estudiantes que responderán las preguntas que hacemos, de modo que a cualquiera le pueda tocar responder en cualquier momento;
b. hacer que varios estudiantes den su respuesta, antes de dar alguna por buena, y pedir a otros estudiantes que opinen sobre las respuestas dadas por sus compañeros;
c. dar tiempo para que los estudiantes piensen su respuesta, evitando la tentación de pasar rápidamente a otro estudiante o ayudarlo con pistas cuando se "atasca"; repreguntar o pedir ampliación en los casos en que las respuestas son muy escuetas, ayudar a que expresen su pensamiento, aunque no esté del todo elaborado;
d. aprender a escuchar y permitir que se expresen ideas erróneas o respuestas equivocadas, justamente para propiciar la discusión; evitar enfocarse únicamente en si la respuesta es adecuada o no. Por el contrario, intentar comprender cuál es el razonamiento del estudiante detrás de su respuesta incorrecta.

Como fue indicado al inicio de este apartado, las preguntas del docente no solo sirven para promover la reflexión. Muchas veces necesitamos tener evidencia rápida que nos permita saber si los estudiantes están comprendiendo o no el punto central en que estamos trabajando, antes de seguir adelante (o para volver sobre el tema con otro abordaje). En palabras de Wiliam (2011: 95), *"es preferible asumir que los estudiantes no han comprendido algo cuando en realidad sí lo han hecho, que asumir que han comprendido cuando no lo han hecho".*

Para este propósito es más útil pensar en preguntas "diagnósticas", que tengan una única respuesta adecuada y, eventualmente, respuestas no correctas que reflejen errores típicos o comprensiones parciales. En el Recuadro 4.4 se incluye un ejemplo propuesto por Wiliam (2011: 98).

Recuadro 4.4: Un ejemplo de pregunta útil para un diagnóstico grupal rápido

¿Por qué los historiadores se preocupan por analizar los problemas de sesgo al analizar testimonios?

a. Nunca se puede confiar en que la gente dice la verdad
b. La gente puede omitir deliberadamente detalles importantes
c. La gente solo puede dar información relevante si ha tenido experiencia directa de los hechos
d. La gente interpreta un mismo evento de maneras diferentes de acuerdo con su experiencia
e. La gente no es consciente de las motivaciones de sus acciones
f. La gente suele confundirse acerca de la secuencia de los sucesos

Fuente: Wiliam, D. (2011: 98).

Esta pregunta fue elaborada por profesores de historia, a partir del análisis de concepciones de los estudiantes y de sus respuestas a distintas preguntas. Una clase en la que todos los estudiantes responden la opción D está lista para seguir adelante. El resto de las opciones son plausibles, tienen algo de cierto, pero no llegan a captar el punto esencial. Si los estudiantes las eligen, eso significa que todavía es necesario discutir y trabajar más el tema.

Un elemento importante al usar este tipo de preguntas con el propósito de obtener un diagnóstico de situación, es poder visualizar rápidamente las respuestas de todos los estudiantes, sin necesidad de hacer una prueba escrita. Para ello las preguntas de opción múltiple suelen ser útiles durante el desarrollo de una clase -más que en pruebas escritas-, empleadas de una forma en que todos los estudiantes respondan en forma simultánea. Una forma práctica de hacerlo (Wiliam 2011: 91) es utilizar tarjetas u hojas, en las que cada estudiante escribe cuál cree que es la opción correcta y todos las muestran al mismo tiempo. Esta técnica es mucho más efectiva que solicitar respuestas orales, porque obliga a todos los estudiantes a pensar por sí mismos en forma simultánea, y le permite al docente contar con un "estado de situación" de todo el grupo, para decidir cómo continuar.

4.3. Proporcionar devoluciones que hacen avanzar el aprendizaje

La tercera estrategia clave de la evaluación formativa es proporcionar a los estudiantes devoluciones que movilicen el aprendizaje. Luego de una extensa revisión de investigaciones, Wiliam aporta pistas muy valiosas para el trabajo en el aula. La primera de ellas está vinculada con un tema al que nos referimos al inicio de este capítulo: la necesidad de reducir el uso de calificaciones. Wiliam (2011: 108-110) refiere dos estudios realizados en Israel por Butler (1987, 1988), en los que se trabajó con varios grupos de estudiantes, cuyos docentes tenían distintas instrucciones en cuanto al tipo de devolución que debían dar.

En el primer estudio, doscientos estudiantes realizaron un mismo conjunto de tareas. Dos días después, a algunos grupos se les devolvieron sus trabajos con calificaciones, a otros grupos con comentarios orientadores, a otros con felicitaciones o estímulos escritos y, finalmente a los restantes sin ningún tipo de devolución. Una semana después los estudiantes debían realizar un conjunto de tareas similares. Los únicos que mejoraron su desempeño fueron aquellos que recibieron comentarios a sus trabajos. Los tres grupos restantes no mostraron ningún avance.

En el segundo estudio se trabajó con 132 estudiantes de doce grupos pertenecientes a cuatro escuelas. Nuevamente, todos resolvieron un mismo conjunto de actividades. Un tercio de ellos recibió su trabajo con una calificación, otro tercio con comentarios, y el tercio restante con ambos, calificaciones y comentarios. En una tarea posterior sobre el mismo tema, los estudiantes que habían recibido solamente comentarios registraron un desempeño superior en un 30% al de quienes habían recibido solamente calificaciones. Pero, lo más interesante de este estudio, fue que quienes recibieron tanto calificaciones como comentarios avanzaron más o menos lo mismo que quienes solo recibieron calificaciones. En otras palabras, si uno va a incluir una calificación en el trabajo de sus estudiantes, casi que no vale la pena destinar tiempo a escribir comentarios.

Esta evidencia es consistente con lo afirmado al inicio del capítulo: el uso excesivo de calificaciones pervierte la finalidad de la enseñanza y desvía la atención de los estudiantes del proceso de aprendizaje. A pesar de ello, tal como mostramos antes, la práctica predominante entre los docentes de primaria y secundaria de la región es asignar puntajes o calificaciones en forma permanente. Muy pocos hacen devoluciones escritas.

Una segunda pista importante es que las devoluciones deben estar enfocadas en la tarea y en aspectos que estén bajo el control del estudiante, es decir, que este pueda modificar. Cuando la devolución se expresa o es percibida como una valoración de la persona (*'qué inteligente'* o *'qué trabajador*

eres', '*te felicito*'), no ayuda a los estudiantes a mejorar, porque está dirigida a rasgos permanentes de su personalidad, no a habilidades específicas. Es más, afirma Wiliam (2011: 111) citando a Brophy (1981): el reconocimiento del docente es efectivo en la medida en que es *"poco frecuente, creíble, contingente (es decir, relativo a un suceso concreto), específico y genuino"*. En otras palabras, la práctica frecuente de muchos docentes, de estimular y felicitar permanentemente a los estudiantes para "elevar la autoestima", con independencia de la calidad del trabajo realizado, no ayuda a mejorar.

Una tercera pista de trabajo es que todos los estudiantes deberían recibir devoluciones y que estas deberían estar orientadas hacia qué hacer a continuación. Muchas veces los docentes solamente le hacemos comentarios a los estudiantes que no han hecho bien su trabajo. Esta práctica tiene dos problemas. Por un lado, los estudiantes enseguida aprenden a interpretar que una devolución de cierta extensión es equivalente a *"no lo hiciste bien"* o *"no me gusta lo que hiciste"*. Por otra parte, la ausencia de devolución para los buenos estudiantes no les permite a estos avanzar en sus aprendizajes (Wiliam, 2011: 129).

Una técnica de trabajo concreta, creada por una docente, es la llamada "de las tres preguntas". Al corregir trabajos, cada vez que esta docente encontraba algo sobre lo que consideraba interesante que el estudiante reflexionase, escribía en ese lugar un número encerrado en un círculo. Esto lo hacía en tres lugares distintos en cada trabajo. Al pie de la hoja o en el reverso, volvía a escribir el número y una pregunta para que el estudiante respondiese. Al inicio de la clase siguiente dedicaba quince minutos para que todos los estudiantes respondieran las preguntas formuladas en sus trabajos. Lo interesante de esta técnica es que todos los estudiantes tenían tarea de revisión para hacer, independientemente de que tan bien o mal estuviese su trabajo (Wiliam 2011: 129).

Una alternativa, para los casos en que se trabaja con ejercicios breves en los que no es fácil hacer comentarios (en matemática, por ejemplo), es evitar marcar cuáles están bien y cuáles mal y hacer una devolución del tipo *"tres de los cinco ejercicios están mal resueltos"*, dejando que el estudiante busque cuáles son, por qué están mal y cómo corregirlos.

Un principio clave, según (Wiliam 2011: 131), es que la devolución debe significar siempre más trabajo para el estudiante que para el docente. Si al docente le lleva más tiempo formular la devolución, que el tiempo que le dedicará el estudiante a trabajar a partir de la misma, algo anda mal.

a. Las devoluciones de un maestro

En el marco de la investigación sobre evaluación en aulas de primaria que realizamos en 2008, encontramos un caso especialmente interesante, un maestro de 6º que fue prácticamente el único de los 160 relevados en 8 países, que regularmente formulaba devoluciones escritas a todos sus estudiantes. Comencemos analizando algunos pasajes de la entrevista en profundidad (Recuadro 4.5). En ellos el maestro explica de qué forma encara la evaluación en el aula con sus alumnos.

Recuadro 4.5: Reflexiones de un maestro sobre sus devoluciones a los estudiantes

..."Uso los deberes para que hagan las correcciones a las devoluciones que les hago. Muchas veces eso no se entiende, porque la tradición escolar dice que el deber debe ser un trabajo diferente. La idea es que en la casa vuelvan sobre lo trabajado en clase..."

"Les devuelvo los trabajos con preguntas. Generalmente resalto alguna cosa que no se tuvo en cuenta. Por ejemplo, el otro día les había dado una propuesta de evaluación con fracciones a las que les faltaba el numerador o el denominador y les pedía que las completaran para que la fracción fuera mayor a 1. Entonces la devolución sería una pregunta como '¿qué pensás sobre esto?'. Es una pregunta para que reflexionen y puedan volver sobre el tema con un objetivo concreto. También podrían ser preguntas que apelen a la explicación que se dio frente a algo. En esos casos deben volver a ver ese trabajo. El problema es el tiempo que demoro para darles los resultados [tiene más de treinta alumnos]".

"Devuelvo con atraso porque me es imposible seguir el día a día si quiero priorizar lo reflexivo. Lo ideal sería que pudiera haber ajuste, de manera que el deber (tarea domiciliaria) sea la continuación de lo que se hace en clase. La idea es que el deber provoque la reflexión que lleve a avanzar".

Consultado sobre la forma de encarar el trabajo con niños que presentan dificultades en su aprendizaje manifiesta:

"...propuestas distintas no hago, esa es una de las salidas para atacar ese problema, sistemas de propuestas distintas. Trato de que los problemas que son planteados puedan tener diversas entradas. Tengo niños que tienen muchas dificultades; en el momento en que ellos hacen distintos trabajos voy pasando por las mesas. Los grupos están organizados de una manera que ninguno pueda apoyarse tanto en el otro que no haga nada, hay distintos niveles porque ningún niño es igual a otro. Pero trato de que cuando llego al lugar donde están los niños con mayores dificultades, uno trata de tirar la punta, la pregunta que pueda ayudar a seguir adelante. Yo preveo posibles procedimientos a resolver. En el caso del niño que quedó en blanco, generalmente tengo alguna idea de cómo preguntarle algo o darle alguna pista para poder hacer algo frente a eso. Una de las herramientas es volver a leer el problema, o ayudarlo a darse cuenta de algo que no tomó en cuenta, otras veces es: por allá dijeron que esto sirve, ¿te parece que puede servir?"

Fuente: Base de registros del estudio **La evaluación en primaria en ocho países de América Latina** (Caso 15).

En las expresiones del maestro se destaca, en primer lugar, su intencionalidad de promover la reflexión de cada alumno sobre su propio trabajo, con el propósito de que avance en el aprendizaje. Esta intención no queda en el plano declarativo, sino que va de la mano con acciones y formas concretas para llevarla a la práctica.

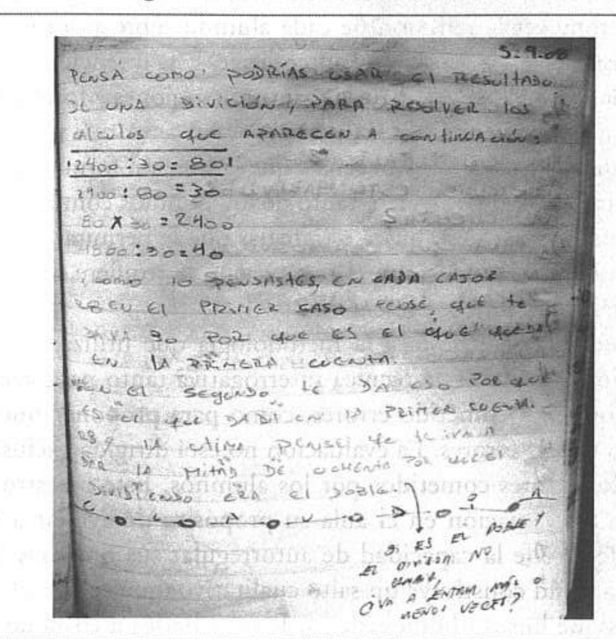
En segundo lugar el maestro señala que la función formativa de la evaluación insume tiempo y esfuerzo, tanto de los alumnos como de sí mismo. Como es consciente de ello, busca estrategias que le permitan armonizar el tiempo realmente disponible con el tiempo que se requiere, sin amilanarse ante los obstáculos que se le presentan.

En tercer lugar se destaca la metodología que utiliza: evaluaciones sin calificaciones y uso de la técnica interrogativa tanto para ayudar a los estudiantes que han cometido errores, como para proponer nuevos retos a quienes no tienen errores. La evaluación no está dirigida exclusivamente a subsanar los errores cometidos por los alumnos. Este maestro pone en el centro de la evaluación en el aula su propósito de ayudar a que cada estudiante desarrolle la capacidad de autorregular sus procesos de aprendizaje. Esta actitud constituye un salto cualitativo, que supera el déficit de los enfoques que buscan homogeneizar los resultados a costa de sacrificar las potencialidades de los estudiantes que no presentan dificultades en sus aprendizajes.

Por último, es interesante destacar que su accionar no es una improvisación frente a cada situación. Si bien comprende que su trabajo exige que dé respuesta a situaciones no previstas (en tanto es imposible tener en cuenta todas las reacciones posibles de los estudiantes en la resolución de una tarea), prevé posibles intervenciones de evaluación formativa desde la planificación.

A continuación analizaremos algunas de las devoluciones escritas realizadas por este maestro a trabajos realizados por sus estudiantes. La Figura 4.3 presenta el trabajo de un niño ante una tarea de matemática sobre el tema división (se incluye transcripción para facilitar la lectura).

Figura 4.3: Devoluciones de un maestro



"Pensá cómo podrías usar el resultado de una división, para resolver los cálculos que aparecen a continuación: 2400:30 = 80"

El alumno realiza equivocadamente la propuesta

4800: 30= 40 y responde:

"te iba dar la mitad de ochenta porque el dividendo era el doble"

El maestro devuelve el trabajo agregándole solamente el siguiente comentario:

"SI ES EL DOBLE Y EL DIVISOR NO CAMBIA, ¿VA A ENTRAR MÁS O MENOS VECES?"

Fuente: Base de registros del estudio **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas.**

En esta tarea es posible observar cómo el maestro efectiviza lo que manifestó en la entrevista: busca promover la reflexión del niño a partir de lo que hizo, partiendo de un error cometido al resolver la tarea. El docente no califica en forma alguna el trabajo analizado. Solamente le plantea al estudiante una pregunta enfocada en la parte del trabajo en que el niño explica erróneamente lo que hizo, por manejar en forma equivocada una propiedad de la división. A partir de esto el docente espera que en un trabajo domiciliario propuesto unos días después de haber realizado la tarea, el alumno reflexione sobre lo hecho a partir de la pregunta planteada, dándole la posibilidad de detenerse con más tiempo a buscar la respuesta correcta para el caso. De esta manera el estudiante podrá buscar en el cuaderno o en el libro, con-

sultar con un compañero o simplemente repensar lo que hizo, para presentar nuevamente el trabajo a su maestro. No se penaliza el error cometido, por el contrario, se ofrece un insumo personalizado para que lo supere el propio estudiante. Por eso consideramos este caso un excelente ejemplo sobre cómo trabajar con los errores desde un enfoque de evaluación formativa.

En la Figura 4.4. presentamos otro trabajo de matemática, del mismo docente. En este caso, el alumno resuelve correctamente la tarea propuesta.

Figura 4.4: Devoluciones de un maestro

Piensen en lo que le convendría saber a alguien para hacer este trabajo más fácilmente. Multiplicar 3 x 20 es fácil. Ahora, ¿cómo se puede usar esa cuenta para hacer 3 x 19?

Expliquen cómo lo pensaron en:

5 x 19 / 7 x 19 / 80 x 19

El estudiante responde:
5 x 20 = 100
7 x 20 = 140
80 x 20 = 1600
100 - 5 = 95
140 - 7 = 133
1600 - 80 = 1520

El maestro devuelve el trabajo con la siguiente pregunta, señalando el 7 de la resta

¿CÓMO SE EXPLICA QUE SE LE SAQUE 7?

Fuente: Base de registros del estudio **La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas.**

Si bien el alumno respondió correctamente, la devolución del docente lo lleva a volver sobre lo realizado, para responder a una pregunta que le exige explicar por qué restó 7 a 140 para resolver 7 x 19, sin realizar la multiplicación. Esto muestra que el docente usa la devolución para promover que el estudiante avance más allá del punto en que se encuentra, en la medida de sus posibilidades. Otro detalle importante es que tampoco en este caso se califica el trabajo, lo cual muestra que el docente no usa calificaciones cuando prioriza la función formativa de la evaluación.

Del análisis de las devoluciones y correcciones que realiza este maestro a los trabajos de sus estudiantes, se constata que sus prácticas en evaluación son consistentes con un enfoque que entiende la función formativa como un proceso que:

- Se desarrolla al interior del proceso de enseñanza e incluye tanto acciones intencionales desde la perspectiva del conocimiento en juego, como espontáneas a partir de situaciones emergentes.
- Su propósito es ayudar a los estudiantes a que avancen por sí mismos desde el punto en que se encuentran. Se propone que cada alumno avance en su reflexión y aprendizaje a partir de lo que ya ha logrado, de acuerdo a sus posibilidades personales.
- Promueve la reflexión del que aprende dándole oportunidades de volver sobre lo realizado en distintos momentos y pudiendo recurrir a otros aportes, además de los del maestro en clase.
- Se enfoca en aspectos concretos del trabajo realizado por el estudiante, sin realizar valoraciones que puedan perturbar el proceso de desarrollo personal del estudiante. Por esta razón no califica los trabajos y lo que hace es ofrecer comentarios directamente vinculados al desempeño en la tarea.
- Tiene un enfoque constructivo del error, utilizándolo para hacer devoluciones productivas a los estudiantes.

b. Las actitudes de los estudiantes ante las devoluciones

Otro aspecto importante a tener en cuenta es que el modo en que cada estudiante recibe las devoluciones depende de muchos factores. Es distinta, por ejemplo, la forma en que recibe una devolución un estudiante que logró un buen desempeño o producto, de la forma en que lo hace otro al que le fue mal. Un estudiante que logró lo esperado podría pensar que el curso es fácil y empezar a esforzarse menos, cuando lo deseable sería que se propusiese metas más ambiciosas (y que el docente contribuya a ello). En cambio, un estudiante que no ha logrado un desempeño o producto satisfactorio, podría

reducir sus aspiraciones, decidir que el curso es demasiado difícil y que no podrá aprobarlo. Por esta razón la cuestión de la motivación es muy importante, según analizamos antes en este capítulo, y no hay recetas: depende del buen juicio y experiencia del docente hacer que sus devoluciones sean apropiadas a cada estudiante.

Los efectos de las devoluciones en cada estudiante dependen especialmente de a qué atribuyen o cómo explican su desempeño (Wiliam 2011: 116-119).

En primer término, es importante notar si el estudiante atribuye su grado de logro en una actividad, a elementos intrínsecos a la misma o a factores externos. Un estudiante que considera que le fue mal porque a su trabajo *"le faltan cosas"*, está haciendo una atribución interna. En cambio, si considera que la devolución recibida obedece a que *"le cae mal al profesor"*, está haciendo una atribución externa. Por lo general los seres humanos tenemos la tendencia a atribuir los éxitos a nuestros propios méritos y los fracasos a factores externos.

En segundo término, hay que observar si el estudiante atribuye su éxito o fracaso a factores permanentes o transitorios. Si el estudiante considera que le fue bien en un trabajo porque él es *"inteligente"* o *"muy capaz"*, está haciendo una atribución permanente. Si considera que le fue bien porque esta vez *"trabajó duro"* o *"se puso las pilas"*, está haciendo una atribución transitoria.

En tercer lugar, es importante observar si el estudiante atribuye su logro o fracaso en una actividad a factores que afectan su desempeño en el conjunto de la disciplina o en el tema específico de la tarea realizada. Dicho en otras palabras, cuando a un estudiante le va mal en una tarea puede pensar que tiene dificultades con ese tema en concreto (por ejemplo, fracciones en Matemática) o puede pensar *"soy malo para las matemáticas"* o, en el peor de los casos, *"soy malo para el estudio"*. En el primer caso está haciendo una atribución específica, mientras que en el segundo hace una atribución global.

A partir del análisis anterior, Wiliam señala que los docentes deberíamos buscar que nuestras devoluciones ayuden a los estudiantes a realizar atribuciones internas -*'las cosas dependen de mí, de lo que yo hago'*-, transitorias -*'mi desempeño en este trabajo depende de cómo lo hice y de cuanto me dediqué'*) y específicas -*'me va mal en algunas actividades pero me puede ir bien en otras, depende del tema'*-. Debemos asegurarnos de que nuestras devoluciones ayudan a los estudiantes a percibir sus capacidades como incrementales y no como fijas, y a pensar en que siempre es posible aprender y mejorar, paso a paso, al propio ritmo. El mensaje principal a trasmitir es *"cómo te vaya en el curso depende de ti y siempre puedes hacer algo para que te vaya mejor"*. Obviamente, no alcanza con decírselo al estudiante, sino que es necesario proveer

orientaciones y oportunidades para rehacer las tareas, de modo que puedan experimentarlo.

4.4. Co-evaluación: activar a los estudiantes como fuente de aprendizaje para sus pares

La cuarta estrategia central de la evaluación formativa consiste en crear espacios y actividades para aprovechar el potencial que los propios estudiantes tienen como fuente de devoluciones relevantes para sus compañeros. Este potencial puede ser promovido a través de actividades de co-evaluación, es decir, actividades en que los propios estudiantes revisan y hacen aportes a los desempeños o trabajos de sus compañeros. Esto forma parte de las denominadas estrategias de "aprendizaje cooperativo", que tienen larga tradición para muchos docentes.

Al mismo tiempo, otros docentes interpretan mal la co-evaluación y la orientan a que los estudiantes califiquen el trabajo de sus compañeros. También suelen interpretar la auto-evaluación como calificarse a sí mismo. Este tipo de aproximación al tema indica insuficiente comprensión de la diferencia entre evaluación formativa y evaluación para la certificación[7].

La co-evaluación tiene muchas ventajas: promueve la cohesión en el grupo, incrementa la motivación de los estudiantes y, algo que no es menor, puede servir para aliviar la sobre carga de trabajo del docente, ya que muchas veces es materialmente imposible dar devoluciones a cada estudiante. Involucrar a los estudiantes en la revisión de los trabajos de sus compañeros enriquece el trabajo en el aula con nuevas visiones. Al mismo tiempo, ayuda a los propios estudiantes a comprender mejor lo que se espera de ellos -desarrollar su "olfato" para las producciones de calidad-. Finalmente, suele ocurrir que los estudiantes captan mejor que el docente las dificultades de sus compañeros o lo que no están entendiendo.

Los procesos de revisión colectiva de trabajos y desempeños ponen en juego habilidades cognitivas y socioemocionales que amplían el campo de acción mental de todos los involucrados. El requisito fundamental es que los criterios de evaluación y las expectativas de logro respecto a la tarea hayan

---

7. La función de certificación le corresponde, en última instancia, al docente, en función de su rol profesional. Esto no significa que el estudiante no pueda opinar sobre la calificación recibida. Asimismo, calificar el propio trabajo puede ser una estrategia formativa, es decir, una manera de hacer que los estudiantes reflexionen sobre sus propios logros. Pero la calificación que se auto asigna el estudiante no debería formar parte de la calificación final.

sido explicadas y comprendidas, al menos parcialmente, por los estudiantes -si bien el proceso de revisión de trabajos de sus compañeros les ayudará a profundizar esta comprensión-. Desde la perspectiva del estudiante cuyo trabajo está siendo revisado, la intervención de sus compañeros le obliga a mirar su propia tarea desde otro ángulo, muchas veces con nueva información, expresada en términos más cercanos a sus posibilidades de comprensión que la ayuda que le puede brindar el propio docente. Si durante este tipo de revisiones aparecen errores en las devoluciones realizadas por los pares, también es una oportunidad de reflexión colectiva, porque se generarán reacciones en otros compañeros. En el caso de que esto no suceda, será una oportunidad para que el propio docente intervenga, haciendo preguntas que obliguen a pensar nuevamente a todos los participantes, generándose así una nueva discusión constructiva.

Por otra parte, este tipo de actividades también promueve el aprendizaje de los estudiantes que brindan ayuda a otros, porque para explicar las cosas deben repensar sus conceptos, mirarlos desde la perspectiva de otro y expresarlos con claridad. Tal como se afirma habitualmente, uno comprende realmente algo cuando es capaz de explicárselo a otro.

Cuatro técnicas para llevar lo anterior a la práctica dentro del aula, tomadas del libro de Wiliam (2011: 137-142), son las siguientes:

a. Cuando los estudiantes trabajan individualmente y alguno tiene dificultades, antes de pedir ayuda al docente debe intentar conseguir esa ayuda de tres compañeros. Recién entonces, si sus compañeros no pudieron ayudarle, el estudiante puede recurrir al docente. Wiliam la denomina *"consulta a tres antes que a mí"*.
b. Una docente de educación media introdujo la práctica regular de que los estudiantes revisaran mutuamente sus tareas domiciliarias, con lo cual redujo su carga de trabajo y logró que cada estudiante recibiera una devolución en forma rápida. En ciertas oportunidades, cada estudiante debe revisar el trabajo de un compañero. En otras ocasiones se hace en grupos de cuatro estudiantes, que deben revisar los trabajos de los integrantes de otro grupo. Los estudiantes que no hicieron la tarea domiciliaria no pueden participar en esta actividad y deben realizar la tarea domiciliaria en el aula. Una de las consecuencias del uso de esta técnica fue un notorio incremento de la cantidad de estudiantes que traían sus tareas hechas. Otra consecuencia notable fue que los estudiantes comenzaron a preocuparse más por la claridad de sus trabajos. Aparentemente les preocupaba más la opinión de sus compañeros que la de su docente.

c. Una técnica complementaria de la anterior es la denominada *"dos estrellas y un deseo"*. Cuando un estudiante revisa el trabajo de un compañero, debe destacar dos aspectos positivos o bien logrados del trabajo (las dos estrellas) y escribir una sugerencia para mejorarlo (el deseo). El autor del trabajo puede luego expresar si está de acuerdo o no con la devolución recibida y si le ha sido de utilidad. El docente recoge luego los trabajos, revisa las devoluciones y, eventualmente, incorpora sus propios comentarios.
d. Otra técnica concreta es denominada por Wiliam *"clasificación de errores"*. En lugar de escribir comentarios en cada trabajo, el docente los devuelve con los errores subrayados. Los estudiantes, trabajando en equipos, deben darse cuenta, caso por caso, en qué consiste el error y de qué tipo es. Por ejemplo, en escritura, los estudiantes deben distinguir entre errores de puntuación, ortografía, concordancia, etc.

4.5. Autoevaluación: activar a cada estudiante como responsable de su propio aprendizaje

Uno de los propósitos principales de la evaluación formativa es lograr que el estudiante se involucre con el aprendizaje y asuma cierto grado de control directo sobre sus formas de aprender. Esto es indispensable porque nadie aprende si no se hace responsable de su propio aprendizaje. De allí la complejidad de la labor educativa y la relación impredecible entre enseñanza y aprendizaje, que analizamos antes en este capítulo.

En su libro Wiliam refiere esta anécdota para ilustrar el punto:

"En la introducción a su libro 'Guitarra', Dan Morgan (1965) escribió, 'Nadie puede enseñarte a tocar la guitarra' (p.1). Esto resultaba bastante extraño, porque el subtítulo del libro es 'El libro que te enseña todo lo que necesitas saber sobre tocar la guitarra'. De todas formas, Morgan aclaró esto agregando [luego de la frase de la primera página], *'Pero te pueden ayudar a aprender'. Esto es bastante obvio. Ya sea aprender a tocar un instrumento musical, jugar un deporte, o una amplia gama de otras actividades humanas, podemos darnos cuenta intuitivamente de que los docentes no crean el aprendizaje; solamente los aprendices pueden crearlo. Y, sin embargo, nuestras aulas parecen basadas en el principio opuesto: que si realmente se esfuerzan, los docentes pueden "hacer" el aprendizaje por los estudiantes. Esto está exacerbado por sistemas de rendición de cuentas que establecen sanciones para docentes, escuelas y distritos, pero no para los estudiantes"* (Wiliam 2011: 145).

De lo anterior se desprende que si pretendemos que nuestros estudiantes aprendan, necesitamos hacer algo para, por un lado, propiciar que comprendan y se apropien de nuestras intenciones educativas y, por otro lado, desarrollar su capacidad para conocer y monitorear sus formas de aprender. Dicho con otras palabras, hay dos elementos centrales para incrementar las posibilidades de aprendizaje de los estudiantes: la motivación y la metacognición. Es crucial que los estudiantes tengan cierto grado de motivación intrínseca por el aprendizaje y, al mismo tiempo, que desarrollen las habilidades metacognitivas necesarias para tomar conciencia de sus fortalezas y debilidades, así como de las formas de mejorar en forma continua y progresiva.

Propiciar estos dos aspectos, motivación y metacognición, requiere, según fue planteado antes, superar la práctica de calificar a los estudiantes continuamente. Asignar notas a todas las tareas de los estudiantes va en contra de los dos elementos mencionados. Por un lado, porque enfoca al estudiante en la motivación extrínseca. Obtener una buena nota pasa a ser su objetivo principal, en tanto que comprender lo que está aprendiendo pasa a segundo plano. En segundo lugar porque las notas ofrecen al estudiante información pobre, demasiado resumida y poco específica sobre su desempeño. Como planteamos antes -y profundizaremos en el capítulo que sigue-, la nota suele ser un promedio agregado que no ayuda a comprender qué es lo que ha logrado, sino qué tanto su trabajo le ha gustado al docente.

a. La motivación

Para analizar la cuestión de la motivación es central la distinción entre motivación intrínseca y motivación extrínseca. Estamos motivados <u>intrínsecamente</u> cuando hacemos algo por la satisfacción interior que nos produce, porque estamos interesados, porque "nos gusta". Disfrutamos haciéndolo, aun cuando nos requiera mucho esfuerzo. Estamos motivados <u>extrínsecamente</u> cuando hacemos algo con el fin de obtener alguna cosa externa, que está más allá de la actividad en sí misma: un premio, una recompensa, dinero, un objeto que queremos poseer (o para cumplir con un requisito para lograr algo más que deseamos). Si queremos conducir un auto, no tenemos más remedio que someternos a un examen para obtener la licencia. Si queremos ingresar a la universidad debemos cursar y aprobar un conjunto de materias que no necesariamente nos interesan. En estos casos nos esforzamos, aún cuando no necesariamente nos guste o disfrutemos de lo que hacemos.

En general los seres humanos necesitamos ambos tipos de motivaciones. Hay muchas cosas importantes que no haríamos si no tuviésemos una exigencia externa, como, por ejemplo, cumplir con un plazo para entregar

una tesis. Al mismo tiempo, solemos hacer mejor aquellas cosas para las cuales tenemos motivación interna. Pero aún teniendo esta última, muchas veces "nos dejamos estar" y no logramos todo lo que podríamos. Paralelamente, muchas veces algo que hacemos para cumplir un requisito externo, termina generándonos motivación y satisfacción interior.

La calificación es una motivación externa, es un requisito para algo que necesitamos: aprobar un curso para poder avanzar en los estudios. Difícilmente un sistema educativo podría sostenerse sin la exigencia externa de demostrar los conocimientos y capacidades adquiridos (de todos modos, es importante notar que esta fuente de motivación externa no opera con estudiantes que no visualizan la educación formal como importante para el futuro). Pero sin motivación interna, el aprendizaje queda en un nivel superficial.

En este punto vale la pena recordar lo que analizamos en el capítulo 2 sobre aprendizaje superficial y aprendizaje profundo. Los estudiantes que logran un nivel de comprensión y aprendizaje profundo son aquellos que están motivados por su propio interés en el tema o actividad en sí misma, no están preocupados por la evaluación externa y tienden a leer más de lo estrictamente necesario. En cambio, los estudiantes con un estilo de aprendizaje superficial, se enfocan en memorizar la información necesaria para la instancia de evaluación como un fin en sí mismo; intentan repetir lo que han "aprendido" tal como aparece en los textos o como ha sido explicado por el docente; y consideran que estudiar significa memorizar datos, hechos, nombres, definiciones y principios. Por lo general se trata de estudiantes que tienen poca autonomía y que se limitan a estudiar lo estrictamente necesario para aprobar. Su motivación principal está relacionada con el miedo a fracasar en la evaluación. Este tipo de estudiantes vive la evaluación casi como una cuestión de suerte, de "acertar" con lo que el profesor espera. No experimentan el aprendizaje como algo que está bajo su control, sino como algo que "les acontece", que no está en sus manos. Por eso, lo único que les preocupa es la nota, como constatación de si han acertado o no en sus respuestas.

Con relación a la motivación intrínseca, Wiliam menciona los trabajos de Csikszentmihalyi (1990), quien ha desarrollado el concepto de "fluir" como el bienestar psicológico resultante de la completa concentración en una tarea, asociado a la sensación de que el tiempo no pasa. El autor tiene varios trabajos publicados en inglés, pero un artículo de Gonzalo Abio (s/f) publicado en castellano sintetiza de la siguiente manera sus principales conceptos:

"...los aspectos aquí mencionados de interés, atención y "no percepción del paso del tiempo" nos llevan a pensar necesariamente en el Modelo de Reto

Óptimo del psicólogo Mihaly Csikszentmihalyi (1975, 1988, 1992), que conduce al sujeto a la llamada "Situación de 'Flujo'", o sea, a un estado en el cual las personas se hallan tan involucradas en la tarea que nada les parece más importante. La situación de flujo constituye una experiencia intrínsecamente placentera que las personas intentarán repetir, siempre que sea posible, para volver a experimentar o sentir esa sensación.

En palabras del propio Csikszentmihalyi (1992, p. 230), las actividades que llevan al ***flujo*** *'son situaciones en las que una persona puede emplear libremente la atención para alcanzar sus objetivos, pues no hay ningún desorden que corregir ni ninguna amenaza para la personalidad de la que haya que defenderse. Hemos llamado a este estado, experiencia de flujo'...*

De forma resumida, diremos que según Csikszentmihalyi, la experiencia de flujo requiere un equilibrio óptimo entre los desafíos percibidos y las habilidades del sujeto. Cuando no existe ese equilibrio, podemos tener dos situaciones diferentes: si los desafíos superan las competencias individuales se genera un estado de ansiedad por exceso de dificultad, mientras que si, por el contrario, las habilidades superan demasiado los desafíos colocados, el individuo estará aburrido y, por ende, poco motivado" (Abio, G. s/f).

De acuerdo con el planteo de Csikszentmihalyi, la motivación del sujeto no depende solamente de la novedad del tema o del interés en el trabajo, sino de la relación entre el desafío que implica la tarea y las posibilidades del individuo para llevarla adelante de manera exitosa. La relación entre ambas dimensiones, desafío y capacidades, puede representarse como en la Figura 4.5, dando lugar a cuatro actitudes básicas hacia la tarea que enfrenta un individuo.

Figura 4.5: La motivación como resultado de la congruencia entre desafíos y capacidades

	Nivel de desafío	
Capacidades	Bajo	Alto
Bajas	**APATÍA**	**ANSIEDAD**
Altas	**ABURRIMIENTO**	**FLUIDEZ**

Fuente: Elaboración propia sobre la base de los aportes de Csikszentmihalyi (1990) y Abio (s/f)

Cuando las capacidades del estudiante son limitadas y la tarea no implica desafío, el resultado es la apatía. Cuando el desafío implicado en la tarea está muy por encima de las capacidades del estudiante, el resultado es

la ansiedad. Por el contrario, si el desafío implicado está muy por debajo de las capacidades, el estudiante se aburrirá.

Obviamente, se trata de una esquematización. Pero lo interesante del planteo de Csikszentmihalyi, afirma Wiliam, es que invierte los términos del análisis de la motivación. Por lo general, nos movemos entre dos extremos. Por un lado, tendemos a percibir la motivación como algo que *'está'* en el cerebro de los estudiantes: tienen motivación, algunos mucha, otros poca. Los culpamos si 'no tienen' suficiente motivación. En el otro extremo, concebimos la motivación como algo que debe producir el docente. Si los estudiantes no aprenden es porque el docente no sabe motivarlos, no es capaz de crear motivación. El culpable en este caso es el docente.

Desde la perspectiva aportada por Csikszentmihalyi, podemos concebir a la motivación no solamente como la causa del aprendizaje, sino también como su consecuencia. Las personas se motivan para aprender algo cuando ven que lo pueden hacer, cuando lo comprenden, cuando le encuentran sentido. Todos los docentes tenemos la experiencia de que los estudiantes se motivan cuando logran entender y avanzar en su aprendizaje, o cuando logran realizar los ejercicios, problemas o tareas que les proponemos. El problema de la motivación, por tanto, se puede enfocar como una cuestión de sintonía entre la dificultad, abstracción y complejidad de los temas y tareas que proponemos en el aula, y las posibilidades de los estudiantes de comprenderlas y llevarlas adelante. Una buena propuesta de enseñanza debería tener esto presente. Y la evaluación debería ayudarnos a afinar la sintonía.

b. La metacognición

Suponiendo que existe motivación, el paso siguiente es ayudar a cada estudiante a que "aprenda a aprender", a que tome conciencia de qué cosas le ayudan a aprender. Esto es la metacognición.

> *"Es el grado de conciencia que tiene una persona acerca de sus formas de pensar (aprender) y de la estructura de sus conocimientos... significa tomar conciencia a partir del análisis evaluativo de:*
> - *cuáles son las formas en que aprende mejor;*
> - *cuándo y por qué aparecen obstáculos y dificultades;*
> - *cómo recuerda mejor;*
> - *cuáles son los dominios de conocimientos que tiene más desarrollados y cuáles menos;*
> - *cuál es el grado de conocimiento que se posee sobre cada estrategia de aprendizaje, así como su uso y aplicación pertinente a situaciones particulares...*

- *...cómo seleccionar una estrategia adecuada para un problema determinado;*
- *cómo determinar si uno comprende lo que está leyendo o escuchando;*
- *cómo transferir los principios o estrategias aprendidos de una situación a otra"* (Celman, S. 1998).

Algunos ejemplos concretos, propuestos por Flavell (1976), son: darse cuenta de que debo chequear o confirmar un resultado antes de aceptarlo como un hecho cierto; saber que es importante revisar con cuidado las cuatro alternativas en una prueba de múltiple opción antes de marcar la que en principio me parece correcta; acostumbrarse a escribir notas o recordatorios sobre ideas importantes que se me ocurren para un trabajo; tener conciencia de la importancia de pedir a alguien su opinión para verificar que mi trabajo está bien encaminado.

A todo lo anterior se podrían agregar algunas habilidades emocionales, como aprender a darse cuenta de los sentimientos que surgen en uno mismo durante el aprendizaje (impotencia, enojo, ansiedad o alegría). Esto es importante para aprender a manejarlos y, sobre todo, para evitar bloqueos emocionales.

Para ayudar al estudiante a desarrollar este tipo de habilidades es necesario que los docentes las tengamos en mente como parte de nuestras intenciones educativas y que desarrollemos actividades específicas para promoverlas. Algunas pistas concretas para hacerlo son las siguientes:

a. **Semáforos**. Es una técnica empleada por muchos docentes para propiciar que los estudiantes desarrollen auto conciencia de su propio aprendizaje. Al inicio de una clase o unidad el docente explica a los estudiantes lo que pretende que aprendan y algunos criterios para valorarlo. Al final de la clase o unidad cada estudiante debe indicar qué tanto cree que ha logrado lo que se esperaba, utilizando los colores del semáforo. Verde si creen que lo han logrado, amarillo si no están seguros o si creen que lo han logrado en forma parcial. Rojo si el estudiante cree que no lo ha logrado. Obviamente, la auto percepción de los estudiantes puede ser errada. Para prevenir esta situación, el paso siguiente consiste en lo siguiente. Los que indicaron rojo se van a trabajar con el docente, mientras que los 'verdes' deben explicar a los 'amarillos' lo que estos no entendieron. Esto permite que el docente y los propios estudiantes confirmen su auto percepción.
b. **Portafolios de progreso**. Otra forma de ayudar a los estudiantes a tomar conciencia de cómo van avanzando progresivamente, consiste en hacerles coleccionar muestras de un mismo tipo de trabajo a lo largo del año. Por ejemplo, dibujos de la figura humana en preescolar o

textos escritos en primaria. Una vez que el estudiante ha realizado varios de estos trabajos a lo largo de varios meses, es notable como ellos mismos pueden identificar aspectos concretos que han ido mejorando. Esto es clave para que tomen conciencia de que su habilidad para aprender es incremental y no fija. Si bien es razonablemente fácil de realizar con cierto tipo de trabajos, puede resultar más difícil cuando se trabaja con ejercicios o actividades breves. El desafío para el docente en estos casos es construir una gama de actividades que refleje una progresión en el tipo de comprensión y habilidades que requieren de sus estudiantes. Los estudiantes pueden ir optando por realizar distintos ejercicios a lo largo de varias semanas o meses, comenzando cada uno por el nivel de dificultad en el que se sienta más cómodo.

c. **Bitácora de aprendizaje.** Es una técnica de uso creciente entre los docentes, que consiste en pedir a los estudiantes que al final de cada clase o unidad registren por escrito algo que han aprendido o que les haya llamado especialmente la atención. Este tipo de registros se pueden realizar tanto en un cuaderno como en plataformas virtuales. Para enriquecer esta práctica Wiliam sugiere que el registro se realice bajo la forma de una lista de frases a completar, de las que cada estudiante debe elegir solamente tres en cada clase. Esto le da más variedad a la actividad y permite que cada estudiante adopte una actitud más reflexiva. Las frases propuestas por Wiliam (2011: 157-158) son:
   - Hoy aprendí...
   - Me sorprendió que...
   - Lo más útil que aprendí en esta clase fue...
   - En esta clase me resultó interesante...
   - Lo que más me gustó de esta clase fue...
   - Una cosa que no estoy seguro de haber entendido es...
   - Lo principal sobre lo que me gustaría saber más es...
   - Después de esta clase me siento...
   - Podría haber aprovechado más esta clase si...

Como resulta obvio, revisando rápidamente este tipo de registros el docente puede aprender bastante sobre su propuesta de enseñanza y sobre el proceso de cada estudiante.

Un comentario final sobre la importancia de trabajar los aspectos metacognitivos. En los últimos tiempos se insiste mucho sobre la cuestión de la diversidad de los estudiantes y sus "estilos de aprendizaje". Muchas veces se afirma en forma un tanto rápida que los docentes deberíamos brindar enseñanza adecuada a las necesidades específicas y al estilo de <u>cada</u> estudiante.

Esto suele ser inviable en la práctica, en virtud de la cantidad de estudiantes con los que se trabaja. En la educación media el profesor muchas veces ni siquiera llega a conocerlos medianamente a todos. Por otra parte, tampoco es bueno que un estudiante desarrolle exclusivamente el tipo de habilidades que le resultan más fáciles. Si a un estudiante le cuesta expresarse por escrito, de allí no se deriva que no deba desarrollar esa capacidad. Lo que los docentes podemos hacer es ampliar la diversidad de propuestas de trabajo y fomentar la co evaluación y la auto evaluación, con el fin de que cada estudiante pueda conocer mejor las cosas que le ayudan a aprender, fortalecerse en su propio estilo, pero también explorar otros estilos y habilidades.

5. Las "rúbricas" como herramienta de evaluación formativa

Queremos cerrar este capítulo presentando una herramienta que consideramos de gran importancia para la evaluación formativa. Se trata de las denominadas "rúbricas", por el término que se utiliza en el mundo anglosajón para designarlas. Las rúbricas son tablas de doble entrada, en las que cada línea incluye un aspecto o dimensión relevante de las intenciones educativas del docente (lo que espera del desempeño del estudiante o del trabajo que está realizando) y en las que cada columna (que por lo general son entre 3 y 5) corresponde a un "nivel de logro" para cada dimensión. El cruce de ambas, dimensiones y niveles, define celdas. En cada celda se debe incluir una descripción de las principales características de un producto o desempeño, en el nivel de logro correspondiente, que muestren una progresión desde los niveles inferiores a los superiores (o viceversa). Al contenido de estas celdas lo denominaremos "descriptores".

En la Figura 4.6 se presenta un ejemplo de rúbrica "vacía". En este caso las dimensiones o criterios serían tres (que habría que definir), y los niveles de logro del producto o desempeño serían cuatro.

Figura 4.6: Ejemplo de rúbrica vacía

Dimensiones o aspectos	Niveles de logro			
	I. En proceso	II. Aceptable	III. Logrado	IV. Destacado
A				
B				
C				

Fuente: Elaboración propia.

La denominación de los niveles depende del enfoque y decisión del docente. En este caso se proponen las siguientes, a modo de ejemplo: *'en proceso'*, para designar un trabajo o desempeño incipiente, que aún no alcanza el nivel de suficiencia; *'aceptable'*, que puede servir para caracterizar lo que se espera como suficiencia mínima; *'logrado'*, que puede incluir las características propias del desempeño o producto que satisface plenamente las intenciones educativas; y *'destacado'*, que caracterizaría los desempeños o productos excelentes, que van más allá de lo esperado. Las formas de designar los niveles pueden variar, según el docente, la disciplina y el tipo de actividad de evaluación. Se las puede designar con términos más cercanos a las calificaciones (por ejemplo, insuficiente, suficiente, muy bueno, excelente) o simplemente con letras o números, para indicar progresión.

5.1. Ejemplos de rúbricas elaboradas por docentes

A los efectos de explicar mejor qué es una rúbrica utilizaremos dos ejemplos concretos de rúbricas elaboradas por docentes en el marco de actividades de formación sobre el tema[8]. El primer caso corresponde a una actividad auténtica en Geografía, elaborada por una maestra de 6º de Primaria. La situación planteada a los estudiantes fue la construcción de un nuevo centro comercial en la ciudad de Salto (Uruguay). Se trataba de una situación real, que había generado una importante discusión pública. Los estudiantes debían elaborar tres argumentos a favor de la construcción del centro comercial, *"teniendo en cuenta las características y funciones de un sistema urbano, su movilidad, expansión y organización"*. Debían además escribir *"qué problemas podría traer para la ciudad en general, la zona de influencia y para las personas que viven en ella, en particular"*.

La rúbrica para la evaluación de esta actividad, que presentamos en la Figura 4.7, además de haber sido elaborada por una maestra en el marco de sus prácticas de trabajo cotidianas, es especialmente interesante por varios aspectos que analizaremos a continuación. Contiene también elementos que se podrían mejorar.

8. El Diploma en Evaluación de Aprendizajes, que se dicta desde el año 2009 en el Instituto de Evaluación Educativa de la Universidad Católica del Uruguay. Estos ejemplos se incluyen en el libro con autorización de sus autoras.

Figura 4.7: Rúbrica para 6º de primaria

DIMENSIONES	PUEDO UTILIZAR LO QUE APRENDÍ	ESTOY EN PROCESO DE APRENDIZAJE	TODAVÍA TENGO DIFICULTADES	CREO QUE DEBO VOLVER A ESTUDIAR EL TEMA
CÓMO ARGUMENTO: • Claridad • Siguiendo la estructura argumentativa • Precisión en el lenguaje	Los argumentos se explican en forma clara, precisa y concreta.	Los argumentos que presento están redactados con coherencia.	Los argumentos son confusos.	Escribo enunciados que no son argumentos.
	Al enunciar cada argumento respeto la estructura argumentativa.	En mis enunciados aparecen por lo menos dos elementos propios de la estructura argumentativa (premisa/ argumentos/conclusión).	En mis enunciados aparecen claramente razones o argumentos.	Mis enunciados no tienen estructura argumentativa.
	Utilizo el lenguaje técnico adecuado al tema.	Utilizo un lenguaje coloquial aunque aparecen palabras propias de la temática.	Utilizo un lenguaje coloquial.	Utilizo un lenguaje coloquial.
ARGUMENTACIÓN TENIENDO EN CUENTA ATRIBUTOS DEL CONCEPTO ESPACIO URBANO: • Alta densidad de población • Actividades económicas predominantes • Organización del espacio	Presento 3 argumentos a favor de la construcción del centro comercial que se relacionan con la alta densidad de población, las actividades económicas de la zona y la forma en que se organiza el espacio.	Presento dos argumentos a favor de la construcción del nuevo centro comercial relacionados con los atributos del espacio urbano.	Presento un argumento a favor de la construcción del centro comercial relacionado con algún atributo del espacio urbano.	Mis argumentos no se basan en las características del espacio urbano, sino en otros factores. .
DESCRIBO POSIBLES PROBLEMAS RELACIONADOS CON LA CONSTRUCCIÓN DEL HIPERMERCADO	Describo tres posibles problemas relacionados con la construcción del nuevo centro comercial utilizando los atributos del concepto espacio urbano.	Describo dos posibles problemas relacionados con el nuevo centro comercial utilizando los atributos del concepto espacio urbano.	Describo un problema respecto al nuevo centro comercial, que se relacione con alguno de los atributos del espacio urbano.	No encuentro problemas para el establecimiento de un nuevo centro comercial en la ciudad.

Fuente: Maestra Cecilia Pereira (2012).

Un primer aspecto a destacar es la forma en que plantea los "niveles" de logro: *"puedo utilizar lo que aprendí"*, *"estoy en proceso de aprendizaje"*, *"todavía tengo dificultades"*, *"creo que debo volver a estudiar el tema"*. Es interesante notar que no tienen carácter valorativo, sino que están formulados como

una auto evaluación por parte del estudiante y expresan su autopercepción con relación al uso del conocimiento estudiado para resolver la tarea. En este sentido, propicia la metacognición. También es interesante la forma de definir las dimensiones, distinguiendo entre la forma de argumentar (técnica argumentativa), por un lado, y el uso de los conceptos de espacio urbano trabajados en clase, por otro. Al mismo tiempo, las dimensiones orientan al estudiante a pensar en argumentos a favor y en contra de la construcción del centro comercial.

En cuanto al contenido de cada una de las celdas de la rúbrica, se puede observar que son claramente descriptivas, no valorativas. Este es un elemento clave de una buena rúbrica. Las celdas deben incluir una descripción de las características de un producto o desempeño bueno, regular o malo, pero no deben ser valorativas en sí mismas. Muchas veces tendemos a incluir juicios de valor en lugar de descripciones.

Un aspecto que se podría mejorar es el uso de cantidad de argumentos como criterio diferenciador entre los distintos niveles. En realidad lo importante no es que haya exactamente 2 o 3 argumentos, sino que haya varios, a favor y en contra, y que los argumentos sean pertinentes, estén bien formulados y estén fundados en los atributos conceptuales estudiados. El uso de cantidades puede inducir a los estudiantes a "cumplir" con incluir una cierta cantidad de elementos en sus trabajos, sin preocuparse por su calidad o pertinencia. Más allá de estas y otras posibles mejoras (por definición una rúbrica siempre es mejorable), se trata de un ejemplo de muy buena calidad.

En el Recuadro 4.6 incluimos parte de las reflexiones que la autora de esta rúbrica realizó en torno a su experiencia con el uso de la misma. Estas reflexiones fueron parte del trabajo de acreditación del Módulo sobre el tema en el Diploma antes indicado. Creemos que las mismas explican mejor que muchos textos teóricos los desafíos implicados en la construcción y uso de rúbricas.

Recuadro 4.6: Reflexiones sobre la construcción y uso de la rúbrica

Luego de presentada la propuesta, y antes de que ellos la realizaran se les solicitó que verbalizaran lo que se les pedía y que pensaran cómo podrían determinar si la actividad era realizada con éxito. Es así que ellos esbozaron un primer borrador de la rúbrica, el cual incluía algunas dimensiones y criterios como, por ejemplo, que los argumentos se enunciaran en forma clara, coherente y con un lenguaje acorde al tema tratado. Cuando les hice notar que no servía cualquier tipo de argumento, volvieron a leer la consigna propuesta y plantearon ejemplos de cómo debía

ser el contenido de esos argumentos y ejemplos de como no debía ser ese contenido, logrando acordar que los argumentos tenían que contener características del espacio urbano. Tomando como punto de partida las ideas de los alumnos acerca de las dimensiones o indicadores a evaluar es que armé una rúbrica que fuera fiel a lo que ellos pensaban que debía evaluar, pero que a su vez cubriera las dimensiones que yo consideraba relevantes.

En una segunda instancia, les expliqué a mis alumnos que cada uno de ellos tendría que evaluar el trabajo de un compañero. Previamente les presenté la rúbrica y la leímos en forma colectiva. Ellos vieron que figuraba en la lista cada indicador que había sido determinado como necesario a la hora de saber si la tarea estaba bien hecha; pero además analizamos los descriptores de cada nivel de logro, explicando o aclarando su significado con ejemplos y contraejemplos. Luego se repartió un trabajo a cada alumno para que lo evaluara de acuerdo a esa rúbrica.

La dificultad que observé es que no todos estaban capacitados para evaluar cualquier trabajo. Era necesario emparejar los niveles ya que los alumnos más competentes en el uso del concepto podían utilizar la rúbrica con mayor facilidad y evaluar cualquier trabajo, pero no pasó lo mismo con los alumnos con menor nivel de comprensión de los conceptos evaluados.

Una segunda dificultad fue que no sabían qué hacer, si marcar con una cruz o, además, justificar dicha opción. Una vez ajustados esos detalles los chicos trabajaron en su mayoría evaluando con mucho acierto el trabajo de sus compañeros.

Por último, en una nueva instancia se devolvió el trabajo con la evaluación de los compañeros para corregir. Al principio, rezongaron porque decían que esa evaluación no les explicaba nada. Les solicité entonces que leyeran cada enunciado de la rúbrica marcado y que compararan con su trabajo para ver si el mismo era una descripción cierta de lo realizado por ellos. Cuando la respuesta era afirmativa, yo les preguntaba qué otra cosa necesitaban para mejorar su tarea. En algunos casos bastó para que mejoraran los enunciados tanto desde el punto de vista de la estructura argumentativa como desde su contenido, ya que se esforzaban por encontrar relaciones entre atributos del concepto espacio urbano y la situación planteada. En otros casos, fue necesario orientar la tarea con ideas más concretas dadas por mí...

...Encontrar niveles de logro para cada dimensión me obligó a pensar en qué posibles indicios iba a encontrar en los trabajos de los alumnos. Sin embargo todas mis ideas previas, las cuales de algún modo se ven reflejadas en la rúbrica que elaboré, se vieron superadas por los trabajos de mis alumnos. Luego de leer sus respuestas, de analizar los argumentos, de tratar de entender cómo realizaban sus operaciones mentales y cuáles eran sus procesos cognitivos, pude entender por qué es tan importante tener trabajos modélicos con los cuales poder ir definiendo cuáles son los logros y las dificultades en cada nivel de logro propuesto.

Fuente: Maestra Cecilia Pereira (2012)

El segundo ejemplo de rúbrica fue elaborada por una profesora de Expresión Artística de Educación Media. La actividad propuesta por la docente a sus estudiantes se presenta en Figura 4.8, y la rúbrica para valorar la actividad en la Figura 4.9. Según se puede apreciar, la actividad corresponde a la etapa final de un proceso de trabajo en torno al Expresionismo. Los estudiantes deben realizar una pintura, seleccionando la técnica a emplear y el soporte. El trabajo artístico debe estar acompañado de una reflexión escrita sobre la obra realizada. La rúbrica fue utilizada como instrumento para un trabajo de co evaluación de las obras y escritos realizados.

Figura 4.8: Consigna de trabajo para los estudiantes

"EN BUSCA DE LOS ROSTROS EXPRESIONISTAS DE MI IMAGINACIÓN"



"La Mirada Roja"
Arnold Schonberg
1910
Expresionismo
Óleo sobre Cartón
32.2 x 24.6cm
Munich
Galería Estatal de Lenbachhaus
Tomada de
Wolf, N. (2004). Expresionismo. Taschen.Alemania

FASE "CUATRO": PRODUCTO FINAL

Realización de una pintura expresionista, trabajando con técnica y soporte a elección del alumno.

El trabajo tendrá establecidas **tres fechas:**
4 de Setiembre / 11 de Setiembre / 25 de Setiembre

Se tomará como referente alguna de las imágenes seleccionadas para representar los sentimientos, sensaciones o emociones de la actividad anterior.

Tamaño mínimo de las experimentaciones: 1/8 Watman
Soporte: hoja canson texturada, caballito liso o lienzo.
Técnica: Oleo pastel, pastel tiza, témperas u Oleos.

La utilización de otras técnicas deberá estar correctamente justificada en función de las características pictóricas del Expresionismo.

La obra deberá ir acompañada de una **REFLEXIÓN** en la cual se argumenten y justifiquen:

- Los sentimientos, sensaciones y emociones que se quiso representar.
- Las características pictóricas de la corriente expresionista.
- Qué artista se ha tomado como referente y por qué se ha seleccionado.
- Y todos los comentarios, apreciaciones y valoraciones que desee aportar al proceso.

ENTREGAR DE PRODUCTO FINAL: MARTES 2 DE OCTUBRE,
AL FINAL DE LA CLASE, LUEGO DE REALIZAR TRABAJO DE CO-EVALUACIÓN

Fuente: Prof. María Anabela Cúneo (2012).

Según se observa en la Figura 4.9, en realidad se trata de dos rúbricas, una para evaluar la pintura realizada y otra para el texto de reflexión sobre la misma. Para cada una de estas partes la profesora utiliza expresiones diferentes para los niveles de logro, adecuados al tipo de desempeño que cada uno requiere (experto, habilidoso, aprendiz y principiante en un caso; profundo, analítico, superficial y limitado en el otro). Con respecto a la obra, la rúbrica valora dos dimensiones principales: el uso de los elementos propios del Expresionismo, en primer lugar, y la comunicación de elementos emocionales y socioculturales por parte del estudiante, por otro.

Figura 4.9: Rúbrica para un curso de arte en Educación Media Superior

		NIVELES				Co-Evaluación	Auto-Evaluación
	DIMENSIONES	EXPERTO	HABILIDOSO	APRENDIZ	PRINCIPIANTE		
OBRA FINAL	Uso de elementos gráfico-plásticos propios del Expresionismo: • Colores contrastantes, fuertes, semiplanos • Adecuación de la técnica	El uso de los colores y los trazos, así como la técnica elegida y su utilización, favorecen el manejo de los elementos gráfico- plásticos propios del Expresionismo.	Se evidencia el uso de colores así como trazos que definen claramente los elementos propios del Expresionismo, aunque la técnica utilizada no se trabajó bien o no era la adecuada para las intenciones expresivas.	El manejo de los elementos gráficos-plásticos no es acorde al estilo expresionista, pero se evidencian aciertos en el manejo de la técnica seleccionada.	El uso de los elementos gráfico plásticos, así como la técnica elegida y su utilización no evidencian las características propias del Expresionismo.		
	Comunicación de situaciones emocionales y socio culturales: • Adaptación de rasgos del rostro • Incorporación de elementos complementarios	La adaptación de los rasgos del rostro y la incorporación de elementos complementarios en la obra, comunican situaciones emocionales y socio culturales.	Si bien se trabajó en la adaptación de los rasgos del rostro y en la incorporación de elementos complementarios en la obra, es ambigua la comunicación de situaciones emocionales y socio culturales.	Las adaptaciones expresivas realizadas en los rasgos del rostro y la incorporación de elementos complementarios aún no logran comunicar situaciones emocionales y socio culturales.	Se crea un rostro que no presenta adaptaciones expresivas en los rasgos del mismo, ni incorpora elementos complementarios que ayuden a que la obra comunique situaciones emocionales y socio culturales.		
		PROFUNDO	ANALÍTICO	SUPERFICIAL	LIMITADO		
REFLEXIÓN	Análisis y fundamentos de la obra realizada: • Representación de sentimientos, sensaciones y emociones • Características pictóricas expresionistas representadas • Artista/s referente/s seleccionado/s	Las argumentaciones se presentan en forma fundamentada y coherente. Sustenta, define y explica sus opiniones desde una perspectiva crítica abarcando otros puntos de vista e incorporando comentarios justificados.	Las argumentaciones presentan explicaciones e ideas personales. Sustenta, define y explica sus opiniones, pero presenta debilidades al considerar y confrontar con otros puntos de vista.	La argumentación es más descriptiva que analítica, incluye algunos aportes personales más allá de lo trabajado en clase sobre el tema.	Se presentan solo algunos aspectos de lo trabajado en clase. El trabajo resulta incompleto y sin argumentación personal.		

Fuente: Prof. Maria Anabela Cuneo (2012).

La graduación de los descriptores para cada nivel está muy bien lograda en cada una de las dimensiones. Los descriptores para el texto reflexivo están especialmente bien logrados. De todas formas, la tarea de plasmar esta graduación en una rúbrica no está exenta de dificultades. Como no sería posible describir todos los posibles detalles de cada pintura, no hay otro camino que recurrir a expresiones que tienen algún grado de generalidad, tales como "la técnica no se trabajó bien", "se evidencian aciertos en el uso de la técnica" o "es ambigua la comunicación de situaciones emocionales y socioculturales". Tal como se señaló antes en este capítulo, muchos de estos aspectos no pueden ser expresados en forma acabada a través del lenguaje formal e implican cierto grado de "olfato" o "intuición" en el estudiante, que solo puede lograrse ilustrando lo expresado en la rúbrica con los productos concretos realizados por ellos. El papel de una rúbrica en estos casos consiste justamente en destacar, poner sobre la mesa, los aspectos que es necesario observar y valorar. Se puede observar también la existencia de dos columnas a la derecha, destinadas a la evaluación de cada producción por el propio autor (autoevaluación) y por un compañero (coevaluación). El modo en que esto fue realizado aparece en las reflexiones de la profesora sobre su experiencia con el uso de esta rúbrica, incluidas en el Recuadro 4.7.

Recuadro 4.7: Reflexiones sobre la construcción y uso de la rúbrica

Valoro positivamente la instancia de utilización de la rúbrica con los estudiantes. Sin lugar a dudas fue necesario realizar adecuaciones prácticas para la instrumentación de la co-evaluación, principalmente dentro del salón y durante el tiempo de realización de la tarea. En clases previas se explicó que parte de la evaluación de la actividad se realizaría a través de la instancia de co-evaluación y a su vez de autoevaluación. Tal vez este haya sido un factor que incidiera positivamente en la seriedad y dedicación con que los alumnos asumieron la tarea. En general el grupo trabajó en forma ordenada y comprometida durante las tres clases en las cuales se desarrolló la actividad. Pero los mayores asombros por parte de los alumnos se observaron al ver en el proceso de construcción lo que cada uno iba logrando hasta llegar al resultado final.

El martes 2 de Octubre, fecha de entrega de la actividad, los alumnos llegaron al salón con sus productos terminados, además de sus bitácoras con los procesos y reflexiones finales. Se les propuso entonces formar parejas y que cada uno intercambiara su trabajo con el de su compañero e inmediatamente se les planteó como tarea que comenzaran a observar los productos detenidamente,

además de leer las reflexiones finales escritas en la bitácora. A continuación, se entregó a cada alumno la rúbrica correspondiente y se explicó que la misma consta de tres dimensiones que serían observadas en la obra final y en las reflexiones de la bitácora. Se leyó cada una de las dimensiones y los niveles de logro que se establecieron y se indicaron las diferencias existentes entre ellos. Por otra parte, se explicó al alumnado que a medida que fueran evaluando el nivel de logro de su compañero podrían ir identificando el nivel de logro al cual consideran que han llegado ellos mismos en su actividad. Por lo cual dejarían registrado en la planilla que se les presentó los resultados de su par y los de ellos mismos. También de antemano se explicó al grupo que esa instancia de la actividad se debía hacer con el mayor respeto hacia el trabajo del compañero. Esto favorecería el intercambio y el desarrollo de una crítica constructiva hacia la tarea lograda.

Al analizar las formas de utilizar la rúbrica por parte de los estudiantes se identificó que varían desde:

a. Algunos realizaron solamente la autoevaluación, pues señalaron que no están aptos para definir el nivel de logro al que ha llegado su compañero (solo dos alumnos no realizaron la co-evaluación).
b. La verificación del cumplimiento de cada uno de los ítems establecidos en las respectivas dimensiones, marcando con la palabra "SI" o "NO", o definiendo con un "TIC" el cumplimiento de los mismos.
c. La realización de explicaciones y argumentaciones dentro de la columna de co-evaluación y autoevaluación, acerca de las dimensiones observadas, argumentando el nivel de logro que asignaron al trabajo realizado por su compañero y al suyo propio.
d. La incorporación de "Flechas hacia arriba" y "Flechas hacia abajo", para indicar el nivel del logro alcanzado dentro de determinado nivel, pero que, por otra parte, definen que no alcanzaron el nivel superior siguiente o nivel inferior anterior.
e. Manejaron la rúbrica de acuerdo a lo esperado (en la columna correspondiente solamente registraron el nivel de logro asignado a cada dimensión).

Esta variedad de formas de usarla evidencian la capacidad de análisis y aprendizaje de los alumnos, principalmente desde su pensamiento y reflexión. Las cinco categorías indicadas muestran diversas vinculaciones con la rúbrica. En

cuanto a la Categoría A, aparentemente los estudiantes no identificaron en los descriptores caminos que les permitieran valorar de manera objetiva el trabajo de sus compañeros. De haber sido así, no sentirían no estar aptos para evaluar.

En cuanto a las categorías B, C y D podemos decir que los alumnos encontraron necesario incorporar otros elementos, como comentarios, flechas y justificaciones. El realizar argumentaciones en las valoraciones de sus pares muestra el respeto hacia el trabajo de su compañero, procurando dejar lo más claro posible el nivel de logro que se estableció al trabajo evaluado. Además, los alumnos necesitaron "dejar su huella" en la rúbrica. Esto evidencia que la misma es un instrumento de evaluación formativa, dado que tuvieron la necesidad de agregar elementos y de esta manera hacerla más propia. Esto permite ver el grado de reflexión y apropiación con que asumieron la tarea. Su intervención en la matriz refleja sus procesos de análisis durante la instancia de evaluación.

Profesora Anabela Cúneo (2012).

Al igual que en el primer caso presentado, las reflexiones de la profesora sobre el uso de la rúbrica son más que elocuentes y no requieren mayores comentarios.

5.2. Sugerencias prácticas para la construcción y uso de rúbricas

El uso de rúbricas es de gran utilidad para llevar adelante la mayoría de las estrategias clave de la evaluación formativa que analizamos a lo largo de este capítulo. En primer lugar, sirven para comunicar y discutir con los estudiantes las intenciones educativas o metas de aprendizaje, de modo de propiciar su comprensión. Esto puede hacerse de diversas maneras, que van desde elaborar la rúbrica con los propios estudiantes hasta entregárselas pronta después de realizada la actividad de evaluación. Depende de la estrategia de enseñanza del docente. Recuérdese en este punto lo señalado antes, en cuanto a que no siempre la comunicación de las intenciones educativas o metas de aprendizaje debe tener lugar antes de las actividades. La discusión previa de la rúbrica puede ser una forma de comunicar y clarificar dichas intenciones. Pero también puede inducir en los estudiantes respuestas mecánicas que se limiten a "cumplir" con los requisitos estipulados. Como fue dicho, todo depende de la estrategia del docente y de la calidad de la rúbrica.

En segundo lugar, una rúbrica puede ser utilizada por el docente como dispositivo de devolución. En vez de escribir comentarios a cada trabajo, el

docente puede entregar a cada estudiante una versión de la rúbrica indicando en qué nivel considera que el trabajo se encuentra en cada dimensión. Esto se puede hacer simplemente entregando la rúbrica, destacando con un resaltador fluorescente o encerrando en un círculo las celdas correspondientes a la evaluación del docente. Se puede luego solicitar al estudiante que exprese por escrito su acuerdo o desacuerdo con la devolución del docente, explicando sus argumentos. Esto lo obliga a reflexionar sobre su trabajo. El docente puede luego dialogar o responder por escrito a los estudiantes que no comprenden o no están de acuerdo con la devolución recibida. Esta es una forma de reemplazar a las notas como herramienta de devolución. Lo interesante del caso es que no se discutirá sobre calificaciones, sino sobre características concretas del trabajo realizado.

En tercer lugar, las rúbricas tienen un papel clave como herramienta para la co evaluación y la auto evaluación: los estudiantes pueden analizar y discutir sus propios trabajos o los de sus compañeros, confrontándolos con una rúbrica. Esta es una nueva posibilidad a incorporar a las pistas de trabajo desarrolladas más arriba en este capítulo.

Las rúbricas se pueden utilizar de distintas maneras y en distintos momentos. A continuación formulamos algunas distinciones relevantes:

a. Es importante distinguir entre *productos* y *desempeños* en las evidencias que se emplean en las evaluaciones. Un producto es alguna forma de trabajo escrito o proyecto. Por ejemplo, un ensayo persuasivo, un proyecto, una obra. El producto tiene alguna forma de existencia permanente, por lo que puede ser analizado varias veces. Un desempeño es el acto de hacer algo por parte del estudiante: una presentación oral, trabajar en equipo, actuar en una obra de teatro, realizar una demostración gimnástica. En estos casos la evaluación tiene que ser realizada en el momento o en forma inmediata, salvo que se utilice el recurso técnico de registrar el desempeño en video. Las rúbricas deben adecuarse al tipo de evidencia que será utilizada en la evaluación.
b. Las rúbricas pueden ser *generales* o *específicas*. Estas últimas están enfocadas en la evaluación de un producto o desempeño concreto. Tal es el caso de los dos ejemplos que vimos más arriba. Son las más usadas y más fáciles de elaborar cuando uno comienza. Las rúbricas *generales* están destinadas a evaluar el resultado de un proceso más largo, por ejemplo, una unidad, un semestre o un curso completo. En este caso la rúbrica expresa las principales dimensiones y niveles de logro que el estudiante debe alcanzar al final del proceso, y suelen incluir menciones a una gama de productos y/o desempeños, que

constituyen la evidencia de lo que el estudiante ha ido aprendiendo y logrando a lo largo de la unidad o curso.

c. Tal como se vio en el segundo caso presentado, muchas veces resulta útil incluir en la rúbrica algún espacio destinado a la justificación, discusión o reflexión acerca de los descriptores seleccionados para caracterizar el producto o desempeño evaluado.

d. Además de constituir una excelente herramienta para la evaluación formativa, las rúbricas pueden servir de base para construir y asignar calificaciones que tengan más sentido para los estudiantes y familias. Sobre este aspecto nos extenderemos en el próximo capítulo.

En cuanto al proceso de construcción de rúbricas, es importante destacar que se trata de un trabajo artesanal y recursivo. Se comienza con borradores o versiones iniciales, se utilizan o se discuten con los estudiantes, a partir de la experiencia se modifican y se van mejorando... no es algo que se haga bien en el primer intento. Su elaboración requiere tiempo y, en lo posible, trabajo conjunto con otros colegas. Pero es importante darse cuenta de que se trata de una inversión de tiempo que luego se recupera haciendo menos ingrata y más rápida la revisión de trabajos de los estudiantes, y que redunda en mayor motivación e involucramiento de estos con las actividades de aprendizaje.

Al inicio es útil mirar ejemplos de rúbricas. Hay muchas, de distintos tipos, calidades y temas, disponibles en sitios de Internet[9], elaboradas para evaluar una enorme diversidad de productos y desempeños: una presentación oral, la participación en el trabajo en equipo, un cuento, un trabajo de ciencias, un proyecto de investigación, por mencionar algunos.

La primera cuestión clave en la construcción de una rúbrica es pensar en términos de lo que se espera que los estudiantes sean capaces de hacer con los "contenidos curriculares" (en sentido amplio) que uno se propone trabajar con ellos. Como ya fue explicado, es importante pensar no solo en términos de enseñanza (temas a "dar" y actividades a realizar) sino, sobre todo, en términos de aprendizaje (qué esperamos que nuestros estudiantes sean capaces de explicar, comprender o hacer al final de una unidad). En el caso de una rúbrica específica para un trabajo de evaluación concreto, por qué proponemos esta actividad, qué esperamos que los estudiantes nos demuestren al realizarla.

De lo anterior se deriva una decisión cuidadosa para definir cuáles serán las dimensiones o aspectos centrales de la rúbrica (ver las Figuras 4.6, 4.7 y 4.9). Deben reflejar aspectos relevantes de las intenciones educativas que se

9. Sugerimos visitar http://rubistar.4teachers.org/index.php.

tengan y deben diferenciarse claramente una de otra. Brookhart (2013: 25) señala que las dimensiones deberían tener las siguientes propiedades:

a. *Apropiadas*: cada dimensión debe reflejar un aspecto relevante de lo que se espera que los estudiantes aprendan.
b. *Definibles*: cada dimensión debe tener un significado claro tanto para el docente como para los estudiantes.
c. *Observables*: cada dimensión debe describir cualidades del desempeño o producto de los estudiantes que puedan ser percibidas por otras personas distintas del evaluado.
d. *Diferenciables*: cada dimensión debe poder distinguirse con claridad de las restantes, es decir, es importante evitar solapamientos entre ellas.
e. *Completas*: el conjunto de todas las dimensiones debe dar cuenta de manera global de lo que se espera que los estudiantes logren en la instancia evaluativa en cuestión o al final de un curso o unidad.
f. *Describibles a través de un continuo*: cada dimensión debería admitir la identificación y descripción de varios niveles progresivos de calidad.

Una segunda cuestión clave para la construcción de rúbricas es pensar en términos de *progresión*. Si bien en el discurso los docentes enfatizamos mucho la importancia del proceso por sobre los resultados, en la práctica nos cuesta mucho concretar qué entendemos por proceso. Necesitamos pensar menos en términos de correcto/incorrecto o de logrado/no logrado, y pensar más en términos de aproximaciones sucesivas a lo que esperamos de nuestros estudiantes. Esto es lo que los "niveles de logro" de una rúbrica permiten visibilizar. En virtud de ello nos ayudan a superar la falsa dicotomía procesos/resultados, a ver los resultados como algo a lo que el estudiante llega en forma progresiva y a preguntarnos por los procesos típicos de cada nivel de desempeño, así como acerca de los procesos a través de los cuáles se puede avanzar de un nivel al siguiente.

La tercera cuestión clave es la escritura del contenido de las celdas de la rúbrica, es decir, los descriptores para cada dimensión y nivel. También en este caso se requiere de mucha experimentación y corrección recursiva. En general se recomienda comenzar por caracterizar dos niveles, el "aceptable" -el mínimo se que espera que todos los estudiantes logren- y el "logrado" o "satisfactorio" -el producto o desempeño que se ajusta plenamente a las expectativas-. Con esos dos mojones iniciales es posible luego definir el resto de los niveles. Es posible definir el "destacado" o "nivel de excelencia" como aquello que excede las expectativas (y que se sabe de antemano que no todos lo alcanzarán). Es recomendable incluir siempre un nivel del tipo "en proceso", para valorar lo que han logrado los estudiantes que aún no llegan

a lo definido como aceptable. En general no conviene usar más de cuatro o cinco, pero tanto la cantidad como la forma de denominarlos depende de cada docente y del tipo de actividad.

Uno de los desafíos más complejos en la construcción de una rúbrica es expresar con claridad las diferencias entre un nivel y el siguiente. Hay que evitar tanto recurrir a elementos excesivamente concretos (como, por ejemplo, la cantidad de argumentos) como a términos puramente valorativos (muy adecuado, excelente, pobre). También es importante que exista cierto paralelismo entre las celdas correspondientes a una misma dimensión, es decir, que ciertos aspectos del trabajo o desempeño se repitan a lo largo de los niveles, mostrando cómo van evolucionando con los mismos.

La cuarta cuestión clave, tal vez una de las más "sofisticadas" del trabajo de construcción de rúbricas y fuertemente vinculada con la esencia de la evaluación, es mantenerse en un plano descriptivo, sin excederse en la concreción ni en la adjetivación. Expresado en términos negativos: una rúbrica no es una lista de control de aspectos visibles que debe cumplir un trabajo, pero tampoco es un conjunto de adjetivos crecientemente positivos, del tipo "pobre", "aceptable", "bueno", "muy bueno". Utilizar una rúbrica siempre requerirá de un cierto grado de inferencia, para pasar de las descripciones incluidas en las celdas a la formulación de una valoración. De allí que cierto grado de subjetividad sea siempre parte de la evaluación, tal como fue explicado antes. No hay forma de sustituir el juicio subjetivo y experto por una lista de control que cualquiera podría utilizar, sin saber nada del tema o la disciplina que está siendo evaluada (incluyendo, por cierto, las disciplinas deportivas y artísticas).

En términos de (Scriven 2013: 31), una rúbrica es *"una definición que conecta un término evaluativo con una clara descripción no evaluativa de lo que está siendo evaluado"*. El "término evaluativo" es el nivel de desempeño. La "descripción no evaluativa" es el contenido de la celda.

Finalmente, algunas sugerencias para el proceso de trabajo en la construcción de rúbricas:

- apoyarse en la revisión de trabajos previos similares realizados por otros estudiantes, que estén en distintos niveles, para luego describir las características de cada uno de ellos en las dimensiones definidas como relevantes;
- resolver uno mismo el problema o actividad y escribir una respuesta similar a la que espera de los estudiantes como plenamente satisfactoria, para luego describir sus características;

- consultar a los estudiantes durante el proceso de redacción, solicitándoles que lean y opinen sobre las versiones iniciales, con el fin de corroborar qué comprenden de lo que estamos escribiendo;
- cuidar el balance en el lenguaje, de modo que respete las exigencias de la disciplina pero sea comprensible para los estudiantes;
- trabajar en forma colaborativa con otros colegas y plantearse la tarea como un emprendimiento acumulativo y de mediano plazo.

6. Resumen y sugerencias para el trabajo en el centro educativo

En este capítulo hemos recorrido un largo itinerario, con el fin de propiciar el desarrollo de prácticas de evaluación formativa. Como profesor de la Universidad de Aukland de Nueva Zelanda, John Hattie realizó durante varios años una extensa revisión de investigaciones, cuyo objetivo fue identificar qué factores tienen impacto en el rendimiento académico de los estudiantes. De acuerdo con este relevamiento, las buenas prácticas de devolución y evaluación formativa son uno de los aspectos de la práctica docente que mayor impacto positivo tienen en el aprendizaje de los estudiantes (Hattie, J. 2008).

Al inicio del capítulo vimos, a partir de evidencia recogida en varios países de la región, que si bien los docentes tenemos un gran aprecio por la evaluación formativa, en la práctica calificamos a los estudiantes la mayor parte del tiempo. La nota es el principal medio que utilizamos para comunicar a los estudiantes nuestra evaluación de sus trabajos y desempeños.

Sin embargo, calificación y evaluación formativa corresponden a dos finalidades completamente distintas de la evaluación y requieren lógicas de trabajo diferentes. Cuando se las confunde en un mismo acto, inevitablemente la calificación predomina por sobre la intención formativa. Las notas se convierten en el objetivo principal de la mayoría de los estudiantes, que comienzan a desarrollar un vínculo superficial con el contenido y las actividades que les proponemos. El aprendizaje pasa a ser más una cuestión de adivinar qué quiere el docente y de hacer algo para satisfacerlo, que un proceso de comprensión, reflexión y producción que está en sus propias manos (las del estudiante).

Para comprender mejor lo esencial de la evaluación formativa, utilizamos una imagen, tres conceptos y cinco estrategias.

La imagen, propuesta por Wiliam (2011), fue la del puente. La evaluación formativa es una acción dirigida a superar la inevitable brecha entre lo que el docente enseña y lo que los estudiantes comprenden y aprenden. Según

fue explicado, se trata de dos procesos diferentes, que no están necesaria ni automáticamente vinculados. Ni siquiera la mejor enseñanza puede garantizar que todos los estudiantes aprendan. En este punto es que la evaluación aparece como un puente que intenta evitar que enseñanza y aprendizajes sean líneas paralelas que nunca se encuentran. La evaluación formativa nos permite identificar las brechas y ayudar a los estudiantes a superarlas, acercándose más a nuestras intenciones educativas. Nos brinda además la oportunidad de revisar, adecuar y mejorar nuestra propia propuesta de enseñanza.

Los tres conceptos clave presentados, siguiendo los trabajos de Grant Wiggins (1998), fueron los de valoración, devolución y orientación. La calificación es esencialmente valoración. Está orientada a establecer y dar cuenta en forma pública el grado en que cada estudiante ha logrado lo que se espera como aprendizaje. Los docentes tendemos a hacer evaluación formativa a través de valoraciones. Sin embargo, lo central de la evaluación formativa son las devoluciones y la orientación. Las devoluciones están constituidas por todas aquellas acciones y dispositivos dirigidos a ayudar al estudiante a reflexionar sobre su producción o desempeño, y a que comprenda la distancia entre lo que ha logrado y lo que se esperaba. La orientación es todo aquello que ayude al estudiante a dar el paso siguiente en la dirección adecuada. En todo caso, volviendo al concepto de valoración que, según planteamos en el capítulo 1 es la esencia de la evaluación, los docentes deberíamos preocuparnos más por lograr que los estudiantes sean capaces de valorar por sí mismos sus propios desempeños y productos.

De acuerdo con el trabajo de Dylan Wiliam (2011), organizamos la práctica de la evaluación formativa en torno a cinco estrategias principales:

- Compartir, clarificar y comprender las intenciones educativas y criterios de logro.
- Diseñar y llevar adelante actividades y tareas que ofrezcan evidencia de lo que cada estudiante está aprendiendo.
- Proporcionar devoluciones que movilicen el aprendizaje en la dirección deseada (que hagan avanzar el aprendizaje).
- Activar a los estudiantes como fuente de aprendizaje para sus pares.
- Activar a cada estudiante como responsable de su propio aprendizaje.

Cada una de estas estrategias fue explicada a lo largo del capítulo, junto con ejemplos y pistas de trabajo tomadas de propuestas elaboradas por docentes en aula.

Sobre el final del capítulo introdujimos una herramienta de enorme potencial y versatilidad para la evaluación formativa: las rúbricas. A través de dos ejemplos concretos de rúbricas, elaboradas y utilizadas en sus clases por

docentes en actividad, explicamos sus principales características, las diferentes formas de utilizarlas y aspectos a tener en cuenta para construirlas.

Sugerencias para el trabajo en el centro educativo

Según se desprende del largo recorrido realizado en este capítulo, las pistas para el trabajo en evaluación formativa son múltiples y diversas. Nuestra sugerencia es trabajar en conjunto con uno o varios colegas de un mismo nivel de enseñanza (y, eventualmente, de una misma área disciplinar), con el fin de "probar" algo nuevo. Les proponemos hacer la experiencia de llevar adelante las siguientes actividades, con carácter exploratorio, a lo largo de un semestre.

1. Desarrollar una estrategia para explicitar y comunicar a los estudiantes las metas de aprendizaje o intenciones educativas de una unidad del curso. Esto implica sentarse delante de los temas y objetivos que el programa plantea para esa unidad y pensar qué se espera que los estudiantes sean capaces de hacer al final de la misma. Puede ser útil recordar la distinción entre productos y desempeños. También puede ser útil volver sobre las grandes categorías de procesos cognitivos trabajadas en el capítulo 2. Es importante intentar escribir en no más de una carilla estas intenciones educativas, de un modo comprensible para los estudiantes y que, al mismo tiempo, respete el lenguaje propio de la disciplina. Conviene hacer un primer borrador y compartirlo con algunos colegas para que nos hagan sugerencias. Una vez escrito este texto breve, es necesario pensar en la estrategia para comunicar las metas o intenciones a los estudiantes. Recuérdese lo expresado en el apartado 4.1 de este capítulo: no se trata de escribir objetivos en el pizarrón o en el cuaderno, sino de desarrollar la intuición de los estudiantes. La comunicación puede hacerse de diversos modos y en distintos momentos, mediante diálogos, una rúbrica, trabajos modelos, etc. Lo importante es que los estudiantes comprendan el sentido de la unidad en la que están trabajando.

2. Elaborar en forma explícita un conjunto de preguntas orientadas a explorar qué es lo que los estudiantes están comprendiendo del tema abordado y el grado en que están avanzando hacia las intenciones educativas definidas. Utilizar estas preguntas durante el desarrollo de las clases de esta unidad. En el trabajo de elaboración de las preguntas, volver sobre las reflexiones y pistas propuestas en el apartado 4.2 de este capítulo.

3. Hacer la experiencia de realizar la devolución de un trabajo o prueba escrita, sin utilizar calificaciones ni hacer mención a las notas. Enfocar el trabajo de devolución en la reflexión de los estudiantes sobre lo que lograron con relación a las metas o propósitos establecidos. Utilizar para ello alguna de las técnicas de trabajo presentadas en el apartado 4.3 de este capítulo.
4. Elaborar en conjunto con otros colegas una rúbrica, siguiendo las orientaciones incluidas en el apartado 5 de este capítulo. Utilizar la rúbrica como instrumento para una actividad de co-evaluación a lo largo de la unidad trabajada en 1). Cada docente integrante del grupo elabora su rúbrica y la confronta con los colegas. Luego cada uno implementa la actividad de co-evaluación en una de sus clases.

En todas las actividades anteriores puede ser de gran utilidad que cada docente participante registre por escrito sus reflexiones sobre la experiencia (lo que "funcionó" y lo que no, lo que aprendió y lo que debería modificar o mejorar), para luego compartirlo con los colegas.

Referencias bibliográficas

Abio, G. (sin fecha). *El modelo de "flujo" de Csikszentmihalyi y su importancia en la enseñanza de lenguas extranjeras*. En: redELE, Revista electrónica de didáctica / español lengua extranjera, Nº 6. Disponible en: http://www.mecd.gob.es/dctm/redele/Material-RedEle/Revista/2006_06/2006_redELE_6_01Abio.pdf?documentId=0901e72b80df9eed. Acceso 26/10/16.

Brookhart, S. (2008). *Feedback that fits*. En Informative Assessment. December 2007/January 2008 | Volume 65 | Number 4:54-59 Disponible en: http://www.ascd.org/publications/educational-leadership/dec07/vol65/num04/Feedback-That-Fits.aspx. Acceso 23/12/16.

Brookhart, S. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria, Virginia: ASCD.

Brophy, J. (1981). *Teacher praise: A functional analysis*. En: Review of Educational Research, 51(1), pp. 5-32. Citado por Wiliam (2011: 111).

Burns, T. y Köster, F. (editores) (2016). *Governing Education in a Complex World*. París: OCDE Publishing, Education Research and Innovation.

Butler, R. (1987). *Task-involving and ego-involving properties of evaluation: Effects of different feedback condition on motivational perceptions, interest and performance*. En: Journal of Educacional Psychology, 79(4), pp. 474-482. Citado por Wiliam (2011: 108-110).

Butler, R. (1988). *Enhancing and undermining intrinsec motivation; the effects of task-involving and ego-involving evaluation on interest and performance*. En: British

Journal of Educacional Psychology, 58, pp. 1-14. Citado por Wiliam (2011: 108-110).

Celman, S. (1998). *¿Es posible mejorar la evaluación y transformarla en herramienta de conocimiento?* En: Camillioni, A. y otros, La evaluación de los aprendizajes en el debate didáctico contemporáneo; Cap. 2, págs. 35-66. Buenos Aires: Paidós.

Chappuis, J., Stiggins, R., Chappuis, S. y Arter, J. (2007). *Classroom Assessment for Student Learning: Doing It Right - Using It Well.* New Jersey: Pearson Education Inc. Disponible en: http://bostonpublicschools.org/cms/lib07/MA01906464/Centricity/Domain/99/stiggins_assessment.pdf. Acceso 23/12/16.

Chappuis, J. (2009). *Seven Strategies of Assessment for Learning.* Assessment Training Institute. Portland: Pearson Education Inc.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience.* New York: Harper & Row.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the Psychology of Discovery and Invention.* New York: Harper Perennial.

Flavell, J. H. (1976). *Metacognitive aspects of problem solving.* En: Resnick, L. B. (ed.), The Nature of Intelligence (pp. 231-235). Hillsdale, New Jersey: Erlbaum. Citado por Wiliam (2011: 148).

Hattie, J. (2008). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement.* New York: Routledge.

Instituto Nacional de Evaluación Educativa (INEEd) (2015). *Evaluación y tránsito educativo. Estudio de propuestas de evaluación en las aulas de educación primaria y media en Uruguay. Resumen ejecutivo.* Montevideo: Instituto Nacional de Evaluación Educativa.

Jorba, J. y San Mart, N. (1993). *La función pedagógica de la evaluación.* Artículo publicado en Aula de Innovación Educativa, Nº 20, pp. 20-30. Disponible en http://www.seg.guanajuato.gob.mx/Ceducativa/CartillaB/6antologia/Referentes%20para%20la%20evaluaci%C3%B3n/Sobre%20estrategias%20y%20herramientas%20para%20llevar%20a%20cabo%20evaluaci%C3%B3n%20f/Jorba,%20Jaume%20y%20S.%20Neus.pdf. Acceso: 23/12/16.

Martínez Rizo, F. (2012). *La evaluación formativa del aprendizaje en el aula en la bibliografía en inglés y francés. Revisión de literatura.* En RMIE, 2012, Vol. 17, Núm. 54, PP. 849-875. Disponible en: http://www.scielo.org.mx/pdf/rmie/v17n54/v17n54a8.pdf. Acceso 23/12/16.

Picaroni, B. (2009). *La evaluación en las aulas de primaria: usos formativos, calificaciones y comunicación con los padres.* Lima: GTEE-PREAL.

Mercado, A. y Martínez Rizo, F. (2014). *Evidencias de prácticas de evaluación de un grupo de profesores de primarias de Nuevo León.* Publicado en la Revista Mexicana de Investigación Educativa, Vol. 19, Nº 61, pp. 537-567. Disponible en: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-66662014000200009. Acceso 23/12/16.

Pirsig, M. (2004). *Zen y el arte de la mantención de la motocicleta*. México D.F.: Sexto Piso.

Ravela, P. (2006). *Fichas didácticas. Para comprender las evaluaciones educativas*. Santiago de Chile: PREAL.

Ravela, P. (2009). *Consignas, devoluciones y calificaciones: los problemas de la evaluación en las aulas de educación primaria en América Latina*. En: Páginas de Educación, Vol. 2, pp. 49-89. Montevideo: Universidad Católica del Uruguay.

Ravela, P., Leymonié. J, Viñas, J., Haretche, C. (2014). *La evaluación en las aulas de secundaria básica en cuatro países de América Latina*. En: Propuesta Educativa, Año 23, Nº 41, pp. 20-45, Junio 2014. Buenos Aires: Facultad Latinoamericana de Ciencias Sociales.

Sadler, D. (1989). *Formative assessment and the design of instructional systems*. En: Instructional Science, Nº 18, pp. 119-144.

Scriven, M. (1967). *The Methodology of Evaluation*. Indiana University. Disponible en: http://www.comp.dit.ie/dgordon/Courses/ILT/ILT0005/TheMethodologyOfEvaluation.pdf. Acceso: 23/12/16.

Shepard, L. (2006). *La evaluación en el aula*. México: Instituto Nacional para la Evaluación de la Educación (INEE), 2008. Traducción del original en inglés, *Classroom Assesment;* en Brennan, R. (ed.) Educational Measurement, 4ta. Edición. ACE/Praeger Westport, pp. 623-46.

Wiggins, G. (1998). *Educative Assessment. Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.

Wiliam, D. (2011). *Embedded Formative Assessment*. Bloomington: Solution Tree Press.

CAPÍTULO 5

EVALUACIÓN PARA LA CALIFICACIÓN Y LA CERTIFICACIÓN

Introducción

En el capítulo anterior explicamos que en el aula tienen lugar dos tipos de evaluación completamente diferentes en su finalidad y su lógica. Una es la evaluación cuyo propósito es informar públicamente acerca del grado en que cada estudiante ha logrado los aprendizajes esperados para el curso que realiza. A esta evaluación la denominamos "de certificación" y está vinculada a la función social de acreditación de conocimientos y capacidades, propia de los sistemas de educación formal. Su función es dar cuenta ante los colegas, las familias, los propios estudiantes, otras instituciones educativas y la sociedad, acerca de lo aprendido por cada estudiante. Cumple además una segunda función importante: establecer una exigencia externa para los estudiantes con relación al aprendizaje. La importancia a otorgar a la evaluación para la certificación debería ser inversamente proporcional al nivel del sistema educativo de que se trate. En los niveles inferiores casi no debería existir (de hecho no existe en la educación inicial), en tanto en los niveles superiores, en los que se otorgan títulos habilitantes para el desempeño de funciones sociales, su importancia es crucial. En primaria debería tener un papel secundario, en tanto que durante la adolescencia su función de exigencia cobra particular importancia como motivación externa.

Por otro lado está la evaluación que solemos denominar "formativa", que tiene una finalidad y una lógica completamente diferentes. Su propósito principal es ayudar al estudiante a avanzar en el aprendizaje, a partir de la constatación de la distancia entre lo que ha alcanzado y las intenciones educativas del docente. La función formativa, concretada a través de las devoluciones y de las consiguientes orientaciones acerca de cómo avanzar, puede ser realizada por el docente, por el propio estudiante y/o por sus compañeros. Además de lo anterior, la evaluación formativa tiene como función aportarle al docente información útil para revisar y ajustar su propuesta de enseñanza.

La evaluación formativa está intrínsecamente ligada a la enseñanza y a los procesos de aprendizaje. Su función es establecer un puente entre la enseñanza y el aprendizaje dado que, como fue explicado, no hay una relación directa y necesaria entre ambas. De allí la necesidad de constatar las brechas que se producen y hacer algo para cerrarlas. La evaluación formativa no está dirigida a informar públicamente acerca de lo aprendido ni a establecer una exigencia externa sino que, por el contrario, su objetivo es propiciar el desarrollo de la motivación intrínseca y lograr que cada estudiante se apropie de su aprendizaje.

Tal como señalamos en el capítulo anterior, la mayoría de los docentes tendemos, en la práctica, a fundir (y confundir) ambos tipos de evaluación (o, en otras palabras, a no distinguir adecuadamente entre ambas). Solemos usar la nota como forma principal de devolución al estudiante y, muchas veces, como instrumento de motivación externa y disciplinamiento. En forma explícita o implícita, calificamos la mayor parte del tiempo todo tipo de actividades y actitudes de los estudiantes. La consecuencia de esta forma de proceder es que la calificación "mata" a la evaluación formativa. La nota se transforma en el principal objetivo del estudiante, que adopta una actitud pasiva y superficial con respecto al aprendizaje. Su motivación pasa a ser complacer al docente, superar a sus compañeros (en el caso de los "buenos "estudiantes) o evitar el fracaso (en el caso de los que tiene dificultades). La relación con el aprendizaje que se genera es superficial y memorística.

Por otra parte, corregir trabajos de los estudiantes es tal vez la parte más ingrata de la tarea docente. Nos requiere una gran cantidad de tiempo y suele ser bastante aburrido. Es una tarea casi administrativa. Algunos la realizamos con mayor cuidado y detalle, otros hacemos una lectura "cruzada" y rápida de los trabajos. Establecer una calificación suele ser relativamente sencillo para los casos extremos (estudiantes muy "buenos" o muy "malos"), pero también resulta estresante (y hasta un poco angustiante) en los casos "dudosos", cuando sabemos que el futuro del estudiante depende en parte de nuestra decisión de aprobarlo o reprobarlo. En estos casos nos debatimos interiormente entre la consideración por el estudiante ("darle una oportunidad") y la necesidad de asegurar que se alcancen ciertos logros o aprendizajes para aprobar, como forma de evitar que el pasaje por el sistema educativo se transforme en una experiencia vacía de contenido ("mantener el nivel"). Consideramos si el estudiante realmente ha puesto todo su esfuerzo, si lo que pretendemos que aprenda está a su alcance y nos debatimos con lo que el programa de estudio nos pide que enseñemos.

A lo largo de este capítulo analizaremos con más detalle los modos en que los docentes calificamos a nuestros estudiantes. Comenzaremos por una

rápida revisión de los marcos normativos que rigen la tarea de calificación (los reglamentos de evaluación), dado que definen un espacio dentro del cual los docentes debemos actuar. Mostraremos las principales prácticas de calificación, tal como emergen de los trabajos de investigación que realizamos en varios países de la región, y sus principales problemas y efectos contraproducentes. Al igual que en capítulos anteriores, propondremos estrategias y pistas de trabajo para mejorar nuestras formas de calificar a los estudiantes.

1. Los reglamentos de evaluación y el eslabón perdido

En el marco de los trabajos de investigación que realizamos entre 2008 y 2012 en nueve países de la región, analizamos los reglamentos de evaluación vigentes en cada sistema educativo, a través de la revisión documental y de las entrevistas con los docentes. El análisis de los reglamentos es importante porque estos definen ciertos elementos clave que encuadran la tarea de calificar que realizamos los docentes: la escala de calificaciones, la frecuencia de las mismas, la "nota de aprobación", los aspectos o criterios a considerar al establecer una calificación. Estos elementos definen un espacio de acción con ciertos límites, pero en modo alguno determinan nuestro accionar. Como veremos, los docentes siempre tenemos, para bien y para mal, cierto margen de "maniobra" dentro de las normas vigentes.

Los reglamentos de evaluación son importantes también porque inciden fuertemente sobre las expectativas y demandas de estudiantes y familias, al tiempo que reflejan la cultura existente en una sociedad en torno al modo de transitar por el sistema educativo y de constatar y reportar el logro de los aprendizajes esperados.

Como los reglamentos de evaluación cambian con cierta frecuencia, en este capítulo trabajamos con las versiones más recientes que encontramos disponibles en los sitios web de cada Ministerio de Educación. Por razones de espacio, optamos por trabajar con siete de los nueve países incluidos en las investigaciones: Colombia, Costa Rica, Chile, El Salvador, México, Perú[1] y Uruguay[2]. La elección de estos países está vinculada a la diversidad cultural

1. El caso de Perú es especialmente interesante porque es el único país que está intentando establecer una vinculación fuerte entre el diseño curricular y el régimen de evaluación, al menos en el plano documental (la implementación es aún incipiente, dado que el diseño curricular vigente fue aprobado en forma reciente).
2. En el caso de Uruguay analizaremos exclusivamente el Reglamento de Evaluación y Pasaje de Grado vigente en la Educación Secundaria Básica (grados 7°, 8° y 9°). Uruguay es un caso particular en la región porque, como resultado de su

y geográfica y, en particular, a la diversidad de sus reglamentos y sus diseños curriculares, así como del grado de descentralización de la evaluación. Intentaremos mostrar que, a pesar de estas diversidades, existen ciertos problemas centrales a los esquemas de calificación que no están resueltos.

En cualquier caso, nuestro propósito no es hacer un estudio exhaustivo, actualizado y crítico de los reglamentos de evaluación vigentes, sino analizar en forma sintética los marcos normativos dentro de los cuales se construyen las prácticas de calificación de los docentes y las actitudes de estudiantes y familias hacia las notas y los "boletines".

La función principal de los reglamentos de evaluación es definir los siguientes cinco aspectos centrales:

a. el tipo de calificaciones (números, letras, frases cortas o juicios, entre otros) y la escala que se utilizará para evaluar a los estudiantes con fines de certificación, junto con la definición de un "punto de corte", que es el valor o categoría de la escala que representa que el estudiante ha logrado lo necesario para aprobar un curso o materia;
b. los aspectos o dimensiones del "desempeño" del estudiante que deben ser objeto de evaluación a los efectos de definir las calificaciones;
c. los criterios y/o procedimientos que los docentes deben seguir para definir la calificación de un estudiante;
d. la frecuencia con que se debe reportar las calificaciones a los estudiantes y sus familias;
e. las reglas que definen la aprobación o reprobación de un grado o nivel educativo.

Por cierto, los reglamentos de evaluación suelen contener una gama de contenidos más amplia. En general, incluyen elementos conceptuales sobre evaluación; gran cantidad de alusiones, desde nuestra perspectiva retóricas, a la evaluación formativa y al hecho de que la misma debe estar al servicio de la "formación integral" del estudiante; orientaciones relativas a acciones de "recuperación" o apoyo a estudiantes que no logran los aprendizajes "esperados"; regímenes de evaluación diferenciados para estudiantes con necesidades o situaciones especiales, entre otras.

---

particular estructura institucional, cada subsistema (Primaria, Secundaria y Educación Técnica) tiene sus propios diseños curriculares y sus propios regímenes de evaluación, elaborados en forma independiente unos de otros. Recién en diciembre de 2016 se ha presentado un primer documento base para una consulta sobre la posibilidad de establecer en "Marco Curricular de Referencia Nacional", por primera vez en la historia de la educación uruguaya.

1.1. Escalas de calificación, reportes y normas para la aprobación de grados

A los efectos de presentar sintéticamente el contenido de los reglamentos y realizar su análisis, hemos organizado la información en dos tablas. La primera (Figura 5.1) incluye los puntos a), d) y e) antes mencionados (los aspectos formales) y la segunda (Figura 5.2) incluye los puntos b) y c) (los aspectos sustantivos). En la Figura 5.1 se indica cuáles fueron los documentos consultados para cada país.

Figura 5.1: Escalas de calificación y criterios de promoción

Escala y punto de corte	Frecuencia y reporte	Requisitos de Promoción
COLOMBIA: Decreto Nº 1290, por el cual se reglamenta la evaluación del aprendizaje y promoción de los estudiantes de los niveles de educación básica y media. Disponible en: http://www.mineducacion.gov.co/1621/articles-187765_archivo_pdf_decreto_1290.pdf. Acceso 16/12/16.		
Art. 5º - Cada establecimiento educativo definirá y adoptará su escala de valoración de los desempeños de los estudiantes. Para facilitar la movilidad de los estudiantes, se deberá expresar la equivalencia con la escala de valoración nacional: Desempeño Superior Desempeño Alto Desempeño Básico Desempeño Bajo Desempeño básico se entiende como la superación de los desempeños necesarios en relación con las áreas obligatorias y fundamentales. Desempeño bajo se entiende como la no superación de los mismos.	LO DEFINE CADA CENTRO EDUCATIVO	LO DEFINE CADA CENTRO EDUCATIVO: Art. 6º - Cada establecimiento educativo determinará los criterios de promoción escolar de acuerdo con el sistema institucional de evaluación de los estudiantes. Asimismo, el establecimiento educativo definirá el porcentaje de asistencia que incida en la promoción del estudiante.

Escala y punto de corte	Frecuencia y reporte	Requisitos de Promoción
COSTA RICA: Reglamento de Evaluación de los Aprendizajes. Decreto Ejecutivo 35355, actualizado el 19/03/14. Disponible en: http://www.pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_texto_completo.aspx?param1=NRTC&nValor1=1&nValor2=65804&nValor3=96343&strTipM=TC. Acceso 16/12/16.		
Art. 6º - La valoración de los aprendizajes, incluyendo la conducta, se realizará según una escala cuantitativa de uno a cien. Los promedios de cada período se consignan sin decimales. Art. 33º - El estudiante que alcanzare un promedio anual igual o superior a 65 tendra condición de aprobado en la respectiva asignatura. Quien no lo alcance tendrá la condición de aplazado.	Art. 21º - El curso lectivo estará dividido en tres períodos escolares, tanto en la Educación General Básica como en la Educación Diversificada.	Art. 35º: A partir del 2º año de Educación Básica, el estudiante que apruebe todas las asignaturas o módulos, tendrá derecho a ubicarse en el año escolar inmediato superior, o bien a ostentar la condición de egresado del respectivo nivel. Art. 36º-37º: El estudiante de I y II Ciclo que fuere aplazado en más de 4 asignaturas tendrá la condición de reprobado, debiendo repetir el año escolar en forma integral. Si fuere aplazado en cuatro o menos asignaturas, tendrá derecho a presentar pruebas de ampliación en las asignaturas aplazadas. A partir del 2º año de EB, el estudiante que una vez realizadas las pruebas de ampliación, hubiese reprobado de forma definitiva alguna de las asignaturas, se considerará reprobado y deberá repetir integralmente el año escolar.
CHILE: Reglamento de Evaluación y Promoción Escolar de Niños y Niñas de Enseñanza Básica. Emitido en 1997 con modificaciones posteriores. Disponible en: https://www.leychile.cl/Navegar?idNorma=71532. Acceso 16/12/16.		
Art. 7º - Los resultados de las evaluaciones, expresados como calificaciones de los alumnos en cada uno de los subsectores, asignaturas o actividades de aprendizaje se anotarán en una escala numérica de 1 a 7, hasta con un decimal. La calificación mínima de aprobación, deberá ser 4,0.	Art 4º - Los alumnos deberán ser evaluados en todas las actividades del plan de estudio, en períodos bimestrales, trimestrales o semestrales (lo define cada establecimiento).	Art. 11º - Serán promovidos los alumnos que: a) hubieren aprobado todas las asignaturas o actividades de su plan de estudio; b) los que no hubieren aprobado una asignatura si su promedio general de logro es de 4,5 o superior, incluido el no aprobado; c) los que no hubieren aprobado dos asignaturas siempre que su promedio general sea de 5,0 o superior, incluidos los no aprobados.

Escala y punto de corte	Frecuencia y reporte	Requisitos de Promoción
EL SALVADOR: Evaluación al Servicio del Aprendizaje y del Desarrollo (MINEDU 2014). Disponible en: disponible en: http://es.slideshare.net/enmaquijada/nuevo-manual-de-evaluacin-de-los-aprendizajes-2014-el-salvador. Acceso 09/11/16.		
Arts. 28º y 36º - Las calificaciones se deben expresar en una escala numérica que va de 1 a 10. Para aprobar un curso o materia el estudiante debe obtener la nota '5,0' en Educación Básica y '6,0' en Educación Media. Dichas notas "representan el logro básico de las competencias de los estudiantes". En el nivel Medio, si la calificación obtenida en una asignatura fuese de 5'5, "se debe aproximar a 6".	Arts. 27º y 35º - En Educación Básica las calificaciones deben ser establecidas con una frecuencia trimestral. Para la Educación Media se establecen 4 períodos de 10 semanas de duración.	Arts. 33º y 40º - A los efectos de lograr la promoción el grado inmediato superior, en el II y III Ciclo de Educación Básica (grados 4º a 9º) y en la Educación Media, los estudiantes deben alcanzar la nota de aprobación en cada una de las asignaturas correspondientes al grado.
MÉXICO: ACUERDO Nº 696 por el que se establecen normas generales para la evaluación, acreditación, promoción y certificación en la educación básica. Publicado por el Diario Oficial de la Federación el 20/09/13. Disponible en: http://basica.sep.gob.mx/uploads/publication/document/526/acuerdo696.pdf. Acceso 16/12/16.		
Art. 8º - En apego a los programas de estudio y con base en las evidencias reunidas durante el proceso educativo, el docente asignará a cada estudiante una calificación en una escala de 5 a 10... Las calificaciones y los promedios que de las evaluaciones se generen, por asignatura, grado escolar o nivel educativo, se expresarán con un número truncado a décimos. Art. 15º - Se tendrán por acreditadas las asignaturas de educación primaria y educación secundaria..., cuando se obtenga un promedio final mínimo de 6,0.	Art. 6º - En el Reporte de Evaluación se incluirán cinco calificaciones parciales y el promedio final de calificaciones por asignatura y grado escolar. Art. 8º - Se debe reportar en cinco bimestres.	Art. 16º - En 2º y 3º de Primaria el alumno será promovido cuando acredite todas las asignaturas, o cuando obtenga un promedio final de 6,0 en el grado, aun si no acredita el total de asignaturas. En 4º a 6º de Primaria y 1º a 3º de Secundaria, "cuando el alumno tenga un promedio final mínimo de 6.0 en cada asignatura, acreditará el grado cursado". En Secundaria el alumno tendrá que repetir 1º o 2º completos cuando al concluir el ciclo escolar, presente cinco o más asignaturas no acreditadas o si al final de los primeros 15 días hábiles del siguiente ciclo escolar, conserva más de 3 asignaturas de 1º y/o 2º no acreditadas. Tendrá que repetir 3º completo cuando al concluir el ciclo escolar, presente 5 o más asignaturas no acreditadas de 1º, 2º y/o 3er grado".

Escala y punto de corte	Frecuencia y reporte	Requisitos de Promoción
PERÚ: Evaluación de los Aprendizajes de los Estudiantes en la Educación Básica Regular. Directiva Nº 004-VMGP-2005. Disponible en: http://www.minedu.gob.pe/normatividad/directivas/DIR-004-2005-VMGP.pdf. Acceso 16/12/16.		
Para Primaria la escala de calificación es literal y descriptiva, con las categorías que siguen (art. 6.2.1). AD = Logro destacado. Cuando el estudiante evidencia el logro de los aprendizajes previstos, demostrando incluso un manejo solvente y muy satisfactorio en todas las tareas propuestas. A = Logro previsto. Cuando el estudiante evidencia el logro de los aprendizajes previstos en el tiempo programado. B = En proceso. Cuando el estudiante está en camino de lograr los aprendizajes previstos, para lo cual requiere acompañamiento durante un tiempo razonable para lograrlo. C = En inicio. Cuando el estudiante está empezando a desarrollar los aprendizajes previstos o evidencia dificultades para el desarrollo de éstos, necesitando mayor tiempo de acompañamiento e intervención del docente de acuerdo a su ritmo y estilo de aprendizaje. En Secundaria la escala de calificación es vigesimal (1 a 20) considerándose a la nota 11 como mínima aprobatoria (art. 6.3.1.1).	Tanto en Primaria como en Secundaria las calificaciones deben ser entregadas en períodos bimestrales o trimestrales (arts. 6.2.2.1 y 6.3.3.1).	La promoción es automática de 1º a 2º grado. En 2°, 3° y 4° los alumnos son promovidos si obtienen mínimo A en la áreas de Comunicación Integral y Lógico-Matemática, y mínimo B en las otras áreas y talleres. Repiten si obtienen C en Comunicación Integral y Lógico-Matemática. En 5° y 6° son promovidos si obtienen mínimo A en las áreas de Comunicación Integral, Lógico-Matemática, Personal Social y Ciencia y Ambiente, y como mínimo B en las otras áreas y talleres. En todos los grados los alumnos repiten si obtienen C en Comunicación Integral y Lógico-Matemática o si no alcanzan los requisitos indicados luego de un Programa de Recuperación. En Secundaria el estudiante es promovido al grado inmediato superior cuando aprueba todas las áreas curriculares. También aquellos que al terminar el programa de Recuperación Pedagógica desaprobaron como máximo un área curricular. Los estudiantes repiten el grado si al término del año escolar desaprobaron 4 o más áreas curriculares.

Escala y punto de corte	Frecuencia y reporte	Requisitos de Promoción
URUGUAY: Reglamento de Evaluación y Pasaje de Grado de Ciclo Básico de Plan de Reformulación 2006. Disponible en: http://www.ces.edu.uy/ces/images/2016/circulares/2/2956.pdf. Acceso 16/12/16.		
Art. 29º - "El Profesor sintetizará la evaluación del proceso en un juicio conceptual personalizado, referido a los aspectos antedichos, y su traducción numérica en la escala del 1 al 12". Art. 57º - "En todas las instancias evaluativas que corresponda, para calificar al alumno se utilizará una escala numérica de 1 a 12, en la cual el valor 6 corresponderá al mínimo de suficiencia".	Art. 62º - Se debe entregar a los estudiantes boletín con calificaciones en cinco momentos del año: a) primera quincena de mayo; b) junio; c) agosto; d) setiembre; e) diciembre, con el resultado final del Curso.	Art. 75º - Serán promovidos los alumnos que: a) tengan calificación de suficiencia en todas las asignaturas (Promoción Total); b) tengan hasta 3 asignaturas del curso con calificación insuficiente (Promoción Parcial). En el período de diciembre inmediato el estudiante deberá rendir examen en dichas asignaturas. Art. 79º - El estudiante deberá repetir el curso cuando mantenga calificación insuficiente en más de la mitad de las asignaturas del curso; o cuando registre más de 25 inasistencias y tenga más de tres asignaturas con calificación insuficiente.

Fuente: Elaboración propia.

La Figura 5.1 permite apreciar diversidad de formas de expresar las calificaciones. En la mayor parte de los países se emplean escalas numéricas, con diversos rangos de valores. En todos los casos se establece un puntaje o categoría que refleja el logro de los aprendizajes necesarios para aprobar un curso. La escala va de 1 a 10 en El Salvador, siendo 5 la nota de aprobación para Educación Básica y 6 para Educación Media. En Uruguay (en educación secundaria) la escala va de 1 a 12, con 6 como calificación que denota "suficiencia". En Costa Rica la escala va de 1 a 100, siendo 65 el puntaje necesario para aprobar. En Chile la escala va de 1 a 7 y la nota de aprobación es 4. México tiene la particularidad de una escala que va de 5 a 10, siendo 6 la nota de aprobación. En la secundaria de Perú la escala va de 1 a 20 y la nota de aprobación es 11. En Colombia y en la primaria de Perú, en lugar de números, se emplean sistemas de cuatro categorías que reflejan diversos grados de dominio de los aprendizajes esperados. En el caso de Colombia la categoría que significa aprobación es "Desempeño Básico", con una sola categoría por debajo ("Desempeño Bajo"). En Perú en cambio, la categoría que significa aprobación es "Logro Previsto", y deja dos categorías para quienes no logran lo esperado ("En proceso" y "En inicio").

Los casos de Chile y Colombia tienen la particularidad de que dejan la definición de las pautas y reglamentos de evaluación en manos de cada establecimiento, incluyendo la frecuencia para el reporte de las calificaciones. Colombia descentraliza incluso la definición de la escala de calificaciones, pero luego solicita que la misma se ponga en relación con las cuatro categorías de la escala nacional[3].

En todos los países analizados las calificaciones deben ser reportadas a los estudiantes y familias con una frecuencia que oscila entre tres y cinco veces al año.

Como parte de la lógica que rige el modelo escolar predominante, para lograr la aprobación de un año lectivo se requiere de la aprobación de la totalidad de los cursos o asignaturas correspondientes al grado en cuestión. El supuesto es que todos los estudiantes deben alcanzar los mismos logros en los tiempos establecidos y en todas las materias[4]. No hay demasiado espacio para la diversidad de intereses o para ritmos distintos, que podrían implicar, por ejemplo, que ciertos logros en Matemática algunos estudiantes los alcancen a los 8 años de edad y otros a los 10. Esta regla tiene algunos matices. En Chile se permite que el estudiante repruebe una o dos de las asignaturas, siempre y cuando la nota promedio de todas las asignaturas, incluyendo las no aprobadas, sea de 4,5 en el primer caso (una materia reprobada) y de 5,0 en el segundo (dos materias reprobadas) (recuérdese que 4,0 es la nota de aprobación para cada curso). En México existe un régimen de tolerancia similar únicamente en 2º y 3º de Primaria. El estudiante puede promover el grado si el promedio de todas las materias es igual o superior a 6,0, independientemente de que haya o no aprobado todas las materias. En Primaria en Perú se priorizan ciertas áreas, en las que el estudiante debe llegar a la calificación A, en tanto que en las restantes puede estar en B.

Pero la norma general es que, para aprobar un grado, el estudiante debe lograr una calificación de suficiencia en cada una de las materias. En la educación secundaria o media se establecen plazos para que el estudiante avance con "materias pendientes" de aprobación, según se puede apreciar en la última columna de la Figura 5.1: hasta 3 en El Salvador, hasta 4 en Costa Rica

3. El hecho es por demás llamativo: en la medida en que cada establecimiento define el currículo, su proyecto educativo, lo que evalúa y cómo lo hace, es difícil establecer luego el vínculo entre las escalas de los diferentes establecimientos. Las mismas seguramente tengan significados muy diferentes entre una escuela y otra.

4. México y Costa Rica tienen regímenes de promoción automática en 1º y 2º de Educación Primaria. En Chile la promoción es automática entre 1º y 2º, así como entre 3º y 4º de Educación Básica. En El Salvador no existe reprobación, salvo casos excepcionales, durante el Primer Ciclo de Educación Básica.

y México, hasta la mitad de las asignaturas en Uruguay, hasta tres en Perú. En cada país se establecen distintos mecanismos de "recuperación", apoyo especial y oportunidades para rendir exámenes o evaluaciones en aquellas asignaturas que no hayan sido aprobadas al final del año lectivo.

Sin embargo, una de las consecuencias más duras y carentes de sentido de estas normativas, es que, al final del proceso, el estudiante que reprueba alguna asignatura debe volver a cursar la totalidad del grado, incluidas las asignaturas que había aprobado. Esto es así en Costa Rica, Chile, El Salvador, México, Perú y Uruguay. En Costa Rica, por ejemplo, se establece expresamente que *"A partir del 2º año de la EB, el estudiante que una vez realizadas las pruebas de ampliación, hubiese reprobado de forma definitiva alguna de las asignaturas, se considerará reprobado en ese nivel escolar y deberá repetir integralmente el año escolar"*. En México se establece que *"En Secundaria el alumno tendrá que repetir 1º o 2º completos cuando al concluir el ciclo escolar, presente cinco o más asignaturas no acreditadas o si al final de los primeros 15 días hábiles del siguiente ciclo escolar, conserva más de 3 asignaturas de 1º y/o 2º no acreditadas"*. Y así sucesivamente, existen normas similares en los restantes países, salvo en Colombia, que deja la definición de estos criterios en manos de cada establecimiento.

Obviamente este es un elemento carente de racionalidad, salvo por el hecho de que el modelo escolar predominante en la región asume como axiomas que cada estudiante debe formar parte de un único grupo-clase y que todos deben aprender las mismas materias y en los mismos tiempos[5]. Contribuye a generar fracaso y abandono escolar, y pone a los docentes "contra la pared", porque sus decisiones a la hora de calificar en su asignatura pueden tener consecuencias mucho más fuertes para el estudiante que tiene dificultades en varias materias.

1.2. Procedimientos, dimensiones y criterios para definir las calificaciones

Definidas las escalas de calificaciones, la frecuencia con que deben ser asignadas las "notas" y los criterios para definir la aprobación o reprobación de un grado (Figura 5.1), el siguiente aspecto que nos interesa especialmente analizar es cómo se llega a establecer una calificación y qué aspectos sustantivos se tienen en cuenta. La Figura 5.2 ofrece una visión resumida de estos aspectos.

5. Para una excelente crítica al modelo escolar vigente, su relación con el modelo de desarrollo industrial ya superado y su desajuste a la realidad de la cultura, la sociedad y las formas de conocimiento y producción contemporáneas, véase el libro de Leonardo Ferrer, "Educación Proyectiva".

Figura 5.2: Procedimientos, dimensiones y criterios para definir las calificaciones

Colombia	
LO DEFINE CADA CENTRO EDUCATIVO "El cumplimiento de esta labor se encuentra supeditado a la emisión de unos referentes claros y públicos que orienten no solo la evaluación, sino el proceso de formación integral que debe surtirse dentro del sistema educativo. Es en este orden de ideas que, el gobierno nacional emitió el Decreto 1290, donde claramente se determina que la evaluación y la promoción de los estudiantes tendrán como criterios nacionales los referentes de calidad emitidos por el Ministerio de Educación, que son los estándares básicos de competencias, las orientaciones pedagógicas o generales y los lineamientos curriculares" (Ministerio de Educación Nacional, 2009. Documento Nº 11, Fundamentaciones y orientaciones para la implementación del Decreto 1290 de 2009: 58)	
Costa Rica	
Art. 15º - Es deber de cada docente; a) comunicar por escrito a sus estudiantes, en las primeras sesiones del curso, los procedimientos y criterios, que se seguirán en materia de evaluación de los aprendizajes de los diferentes componentes de la calificación; b) entregar a los estudiantes, por escrito y en forma detallada, los objetivos específicos y contenidos que serán medidos en las pruebas, al menos ocho días antes de la aplicación de las mismas. Art. 22º - De los Componentes de la Calificación. La nota de los estudiantes en cada asignatura se obtendrá sumando los porcentajes correspondientes a las calificaciones obtenidas por los estudiantes en los siguientes componentes: Trabajo cotidiano Trabajo extraclase Pruebas Proyecto (según corresponda) Concepto (el juicio profesional valorativo y global que emite el docente con respecto al desempeño y actitud que demuestra el estudiante durante el proceso de aprendizaje) Asistencia. Art. 30º - Del valor porcentual de cada uno de los componentes de la calificación. La calificación de los aprendizajes del estudiante en cada asignatura, será el resultado de la suma de los siguientes valores porcentuales:	
En el II Ciclo de la Educación General Básica: • Trabajo Cotidiano 30% • Trabajo Extraclase 10% • Pruebas (mínimo dos) 50% • Concepto 5% • Asistencia 5%	En el 8º y 9º año de III Ciclo de la Educación General Básica: • Trabajo Cotidiano 15% • Trabajo Extraclase 10% • Pruebas (mínimo dos) 65% • Concepto 5% • Asistencia 5%

Chile
Art. 2º - El Director del establecimiento a propuesta del Consejo de Profesores establecerá un Reglamento de Evaluación sobre la base de las disposiciones del presente decreto. Este Reglamento deberá ser comunicado oportunamente a todos los alumnos, padres y apoderados, a más tardar en el momento de la matrícula. Art. 3º - El Reglamento de Evaluación de cada establecimiento educacional deberá contener, entre otras: a) Disposiciones respecto a estrategias para evaluar los aprendizajes de los alumnos; b) Formas de calificar y comunicar los resultados a los alumnos, padres y apoderados; c) Procedimientos que aplicará el establecimiento para determinar la situación final de los alumnos. Art. 6º - Al término del año lectivo los establecimientos educacionales que así lo determinen podrán administrar un procedimiento de evaluación final a los alumnos en los subsectores, asignaturas o actividades de aprendizaje que consideren procedente. Lo anterior, sin perjuicio de eximir de esta obligación a aquellos que presenten un logro de objetivos que el establecimiento considere adecuado. En los subsectores, asignaturas o actividades de aprendizaje en que se aplique un procedimiento de evaluación final, éste tendrá una ponderación máxima de un 30%.
El Salvador
Art. 7.3 - La evaluación sumativa permite tomar decisiones de refuerzo educativo, recuperación y promoción. Debe ser desarrollada según la planificación del docente y comunicada a estudiantes, familias o representantes. En esta planificación deben considerarse: indicadores de logro, criterios de evaluación, tiempos y valores porcentuales asignados a las actividades. Art. 17 - Criterios de evaluación. Los criterios de evaluación deben ser coherentes con las competencias e indicadores de logro definidos en los programas de estudio. Deben ser definidos por el docente y dados a conocer a estudiantes, familia o representantes. Arts. 27 y 35 - El procedimiento a seguir para establecer la calificación debe considerar tres tipos de actividades: a) "una o más actividades integradoras que contemplen la resolución de problemas y la aplicación de los tres tipos de contenidos [conceptuales, procedimentales y actitudinales], por cada asignatura", (35% de ponderación); b) "actividades cotidianas, como la revisión de cuadernos, trabajos grupales, tareas, portafolios, entre otros", (35% de ponderación); c) "una o más pruebas" (30% de ponderación). Se establece además, tanto para Educación Básica como para Media, que la autoevaluación y la coevaluación de los estudiantes debe tener un peso del 10% en las actividades de tipo 'a)' (integradoras) y 'b)' (cotidianas) (pero no para las pruebas). Art. 31 - Para establecer la calificación trimestral en cada asignatura, "los promedios de las tres actividades de evaluación se registrarán, de forma numérica, en porcentajes equivalentes a la nota obtenida. Por ejemplo: Si un estudiante, en la actividad 1 (con valor del 35%) obtiene 8, le corresponderá la nota 2.8 sin aproximar". La nota trimestral resulta de la suma de las tres actividades, "aproximándola a números enteros".

México
Art. 3º - "La evaluación se basará en la valoración del desempeño de los alumnos en relación con los aprendizajes esperados y las actitudes que mediante el estudio se favorecen, en congruencia con los enfoques didácticos de los programas de estudio de educación preescolar, primaria y secundaria. Asimismo, la evaluación tomará en cuenta las características de diversidad social, lingüística, cultural, física e intelectual de los alumnos" Art. 8º - "En apego a los programas de estudio y con base en las evidencias reunidas durante el proceso educativo, el docente asignará a cada estudiante una calificación en una escala de 5 a 10". Art. 9º - "En los grados de 3o. de primaria a 3o. de secundaria se aplicará un examen final que servirá para calificar el quinto bimestre" Art. 11º - "Promedio final de grado escolar: Será el resultado de sumar los promedios finales de las asignaturas y dividirlo entre el número total de asignaturas que se establecen para cada grado de la educación primaria y educación secundaria en el plan de estudios de educación básica".
Perú
En Primaria: Art. 6..2.2.1: El calificativo bimestral o trimestral del área o taller curricular resulta de la ponderación de los calificativos de las competencias (criterios). El calificativo de cada competencia se obtiene de la ponderación de los resultados que se evidencian en los indicadores de logro. Art. 6..2.2.2: El calificativo anual del área o taller curricular, considerando que la evaluación es un proceso, es el mismo que obtuvo el estudiante en el Área o Taller en el último período (bimestre o trimestre). En Secundaria: Art. 6.3.2.1: Los criterios de evaluación en la educación secundaria son las capacidades de área. Art. 6.3.2.2: En cada área, las actitudes también constituyen un criterio de evaluación del área. Art. 6.3.3.1: a. Al finalizar cada período, los estudiantes tendrán un calificativo por cada criterio de evaluación. Éste se obtendrá promediando los calificativos que los estudiantes hayan obtenido en dicho criterio durante el período de estudios (bimestre o trimestre). b. En cada período también se obtendrá un calificativo de área que resultará al promediar los calificativos de los criterios de área. Art. 6.3.3.2: El calificativo anual del área o taller se obtendrá al término del año escolar, promediando en forma simple los promedios obtenidos en cada período de estudios (bimestre o trimestre).

Uruguay
Art. 23 - "El profesor planificará el desarrollo de su curso y adecuará los objetivos de la asignatura al nivel de cada grupo, centrados en la evolución deseada de los estudiantes, tomados individualmente y como grupo (...) La evaluación consistirá en verificar el grado en que los estudiantes alcanzaron los objetivos propuestos".
Art. 26 - "La actuación general de cada estudiante se evaluará en relación con los objetivos propuestos por el profesor, tomando en cuenta fundamentalmente, los procesos realizados por el estudiante".
Art. 27 - "La evaluación del proceso incluirá, además de la apropiación de los contenidos propios del curso, la apreciación de la actualización del alumno en relación a los siguientes aspectos: interés, actitud en el trabajo, integración social y asiduidad. Se tendrán en cuenta, además, los progresos educativos, la adquisición de valores y el desarrollo de aptitudes".
Art. 28 - "...Se valorarán tareas orales y escritas (de ejercitación y de creación, trabajos de investigación, pruebas de laboratorio), así como trabajos de equipo, en el aula o en otro ámbito".
Art. 29 - "El Profesor sintetizará la evaluación del proceso en un juicio conceptual personalizado, referido a los aspectos antedichos, y su traducción numérica en la escala del 1 al 12".

Fuente: Elaboración propia.

Del estudio de las normas de evaluación resumidas en la Figura 5.2 surgen cuatro elementos principales que nos interesa destacar.

a. Falta de significado sustantivo de las calificaciones en términos de aprendizaje

El primero de ellos es que los significados sustantivos de las calificaciones en general y, en particular, de la "nota de suficiencia" que define la aprobación de un curso, no están definidos específicamente en términos de aprendizajes a lograr. Dicho en otras palabras, ningún reglamento define claramente qué es un desempeño '5' y que diferencia tiene con un desempeño '10', por ejemplo, en Matemática en 5º grado de Primaria. Salvo en una reciente propuesta curricular en construcción en Perú, que comentaremos enseguida, todos los reglamentos de evaluación hacen alusiones genéricas tales como *"La evaluación se basará en la valoración del desempeño de los alumnos en relación con los aprendizajes esperados y las actitudes que mediante el estudio se favorecen, en congruencia con los enfoques didácticos de los programas de estudio"* (México); *"verificar el grado en que los estudiantes alcanzaron los objetivos propuestos"* (Uruguay); o *"La decisión de promoción se hará cuando un estudiante haya alcanzado las competencias mínimas para el grado o nivel superior al que será promovido"* (El Salvador), lo cual remite al grado de claridad con que

estas competencias mínimas estén expresadas en el currículo y a la capacidad de docentes y escuelas para evaluarlas.

Obviamente, un reglamento de evaluación no es un documento en el que se pueda definir el significado específico de cada calificación para cada grado y materia. Por eso los reglamentos remiten a los programas de estudio. El problema es que, salvo excepciones, los programas de estudio de las asignaturas no expresan de manera clara los niveles de desempeño esperados y su relación con las calificaciones. Como norma general, los programas de estudio y los reglamentos de evaluación son elaborados por separado, por equipos diferentes y no tienen puntos de conexión específicos. Esto es lo que llamamos "el eslabón perdido" entre los diseños curriculares y los reglamentos de evaluación. La única forma de articularlos es que cada programa incluya una clara relación entre los aprendizajes y desempeños esperados, y la escala de calificaciones, especialmente en lo que refiere a la nota de aprobación o "suficiencia". Deberían incluir, además, una amplia diversidad de productos y/o desempeños mediante los cuales los estudiantes puedan demostrar sus aprendizajes. Es importante recordar, en este punto, lo planteado en el capítulo 3 con respecto a la evaluación auténtica y las "exhibiciones de comprensión".

Recuadro 5.1: Las calificaciones en el nuevo diseño curricular de Perú[6]

En este contexto es interesante la incipiente experiencia de desarrollo curricular en curso en Perú, si bien en el momento de escribir este texto el nuevo régimen de evaluación todavía está inacabado.

El documento Currículo Nacional de la Educación Básica, publicado al inicio del año 2016, establece 29 competencias con 8 niveles de desempeño progresivos para cada una. Cada nivel de desempeño explicita lo que se espera logren los estudiantes al final de cada ciclo de enseñanza (estándares).

Los ciclos escolares, de 2 o 3 años de duración, son los siguientes: I: 0-2 años; II: 3 a 5 años; III: 1º y 2º de primaria; IV: 3º y 4º de primaria; V: 5º y 6º de primaria; VI: 1º y 2º de secundaria; VII: 3º, 4º y 5º de secundaria. Lo anterior significa que la descripción de niveles de desempeño en cada competencia cubre la totalidad de los grados escolares en forma longitudinal, desde la educación infantil hasta el final de la educación secundaria. El nivel 8 de cada competencia corresponde a un desempeño destacado, que va más allá de lo esperado para el final de la secundaria.

6. Documento disponible en: http://www.minedu.gob.pe/curriculo/pdf/curriculo-nacional-2016-2.pdf. Acceso 16/12/16

Los ciclos escolares, de 2 o 3 años de duración, son los siguientes: I: 0-2 años; II: 3 a 5 años; III: 1° y 2° de primaria; IV: 3° y 4° de primaria; V: 5° y 6° de primaria; VI: 1° y 2° de secundaria; VII: 3°, 4° y 5° de secundaria. Lo anterior significa que la descripción de niveles de desempeño en cada competencia cubre la totalidad de los grados escolares en forma longitudinal, desde la educación infantil hasta el final de la educación secundaria. El nivel 8 de cada competencia corresponde a un desempeño destacado, que va más allá de lo esperado para el final de la secundaria.

En la página 105 se establece una escala de calificación común a todas las modalidades y niveles de la Educación Básica (incluida la secundaria básica). La escala es similar a la presentada en la Figura 5.1, si bien la formulación tiene algunos matices que la mejoran.

AD - Logro destacado: el estudiante evidencia un nivel superior a lo esperado respecto a la competencia.

A - Logro esperado: el estudiante evidencia el nivel esperado respecto a la competencia, demostrando manejo satisfactorio en todas las tareas propuestas y en el tiempo programado.

B - En proceso: el estudiante está próximo al nivel esperado respecto a la competencia, para lo cual requiere acompañamiento durante un tiempo razonable para lograrlo.

C - En inicio: el estudiante muestra un progreso mínimo en una competencia de acuerdo al nivel esperado. Evidencia con frecuencia dificultades en el desarrollo de las tareas, por lo que necesita mayor tiempo de acompañamiento e intervención del docente.

El nuevo diseño curricular tiene la clara intención de vincular estas categorías de calificación con los niveles de desempeño en las 29 competencias. La calificación con fines de promoción "establece conclusiones descriptivas del nivel de aprendizaje alcanzado por el estudiante, en función de la evidencia recogida en el período a evaluar" y debe vincular dichas conclusiones con la escala de calificación(AD, A, B o C).

"Las conclusiones descriptivas son el resultado de un juicio docente realizado basado en el desempeño demostrado por el estudiante, en las diversas situaciones significativas planteadas por el docente. Dichas conclusiones deben explicar el progreso del estudiante en un período determinado con respecto al nivel esperado de la competencia (estándares de aprendizaje), señalando avances, dificultades y recomendaciones para superarlos. En ese sentido, no son notas aisladas, ni promedios, ni frases sueltas, ni un adjetivo calificativo".

"Es importante que estas conclusiones se hagan a través de docentes con base a evidencia variada y relevante del desempeño del estudiante recopilado durante el periodo de aprendizaje a evaluar. Este análisis debe centrarse en los progresos del aprendizaje de cada estudiante en relación al nivel esperado"(105).

Basado en las conclusiones y en la calificación obtenida, el docente deberá elaborar un "Informe de Progreso" del aprendizaje de cada estudiante, dirigido a ellos y a los padres de familia. Este Informe de Progreso deberá ser "entregado de manera personal (al estudiante y a los padres de familia) con el fin de explicar con mayor detalle el nivel actual del aprendizaje del estudiante respecto del nivel esperado de las competencias (estándares de aprendizaje). Así también debe brindar sugerencias que contribuyan a progresar a niveles más complejos".

El enfoque es innovador, porque se propone establecer una relación explícita entre los logros establecidos en el currículo y las calificaciones. De todas formas, en el momento de escribir este libro se trata todavía de una experiencia nueva, cuya implementación concreta habrá que observar. Por el momento no están definidas las normas e instrumentos concretos de evaluación ni, por cierto, su puesta en práctica en las aulas.

b. Subjetividad y heterogeneidad de significados de las calificaciones

Un segundo elemento importante a destacar es que, como consecuencia del punto anterior, la definición de qué significa cada tramo de la escala de calificaciones queda básicamente en manos de cada docente. Por ejemplo:

> *"Los criterios de evaluación deben ser coherentes con las competencias e indicadores de logro definidos en los programas de estudio. Deben ser definidos por el docente y dados a conocer a estudiantes, familia o representantes"* (El Salvador);
>
> *"Es deber de cada docente... comunicar por escrito a sus estudiantes, en las primeras sesiones del curso, los procedimientos y criterios, que se seguirán en materia de evaluación de los aprendizajes de los diferentes componentes de la calificación"* (Costa Rica);
>
> *"El profesor planificará el desarrollo de su curso y adecuará los objetivos de la asignatura al nivel de cada grupo, centrados en la evolución deseada de los estudiantes, tomados individualmente y como grupo... La evaluación consistirá en verificar el grado en que los estudiantes alcanzaron los objetivos propuestos"* (Uruguay).

Este es, en principio, un criterio saludable. Los docentes somos quienes debemos dar cuenta de lo aprendido por cada estudiante. Asimismo, debemos evaluar de un modo coherente con el tipo de trabajos que realizamos durante nuestro curso. Sin embargo, hay dos problemas que no están suficientemente considerados: el aislamiento que caracteriza el trabajo docente y la insuficiente formación que tenemos para evaluar. Como consecuencia de estos dos hechos, el significado de las calificaciones varía enormemente de un docente a otro. En general los docentes tenemos poco

tiempo, escasos espacios colectivos e insuficiente formación para abordar adecuadamente la tarea de evaluación con fines de certificación. Esto da lugar a una enorme heterogeneidad de enfoques y formas de proceder, que solo puede generar ambigüedad en torno al significado de las calificaciones para estudiantes y familias. Ilustraremos enseguida esta afirmación, a través de la descripción de los enfoques y prácticas de calificación observadas en varios países de la región.

c. La "matematización" de las calificaciones

Un tercer elemento a destacar a partir de la sistematización presentada en la Figura 5.2, es el énfasis que ponen algunos reglamentos en establecer un procedimiento matemático preciso para llegar a la calificación. Nos referimos especialmente a los casos de Costa Rica y El Salvador. En este último se establece que el procedimiento a seguir para establecer la calificación debe considerar tres tipos de actividades:

a. *"una o más **actividades integradoras** que contemplen la resolución de problemas y la aplicación de los tres tipos de contenidos* [conceptuales, procedimentales y actitudinales], *por cada asignatura"*, que tiene un peso del 35% en la calificación;
b. ***actividades cotidianas**, como la revisión de cuadernos, trabajos grupales, tareas, portafolios, entre otros"*, que también tiene un peso del 35% en la calificación;
c. *"una o más **pruebas**"*, con un peso del 30%.

Para establecer la calificación trimestral en cada asignatura, *"los promedios de las tres actividades de evaluación se registrarán, de forma numérica, en porcentajes equivalentes a la nota obtenida. Por ejemplo: Si un estudiante, en la actividad 1 (con valor del 35%) obtiene 8, le corresponderá la nota 2.8 sin aproximar"* (MINEDU, 2014: 60). La nota trimestral resulta de la suma de las tres actividades, *"aproximándola a números enteros"*. Se establece además, tanto para Educación Básica como para Media, que la autoevaluación y la coevaluación de los estudiantes debe tener un peso del 10% en las actividades de tipo 'a)' (**integradoras**) y 'b)' (**cotidianas**) (pero no para las **pruebas**).

En el caso de Costa Rica el reglamento de evaluación dedica trece páginas, con numerales que van de la a) a la w), a establecer diferentes fórmulas de ponderación para llegar a la calificación en las distintas asignaturas, niveles y modalidades educativas.

El aspecto positivo de este tipo de normas es que propician que los docentes empleen diversidad de fuentes de información a la hora de evaluar. El aspecto negativo es que inducen a los docentes a construir la calificación

como una suma y ponderación de puntajes, perdiendo de vista los desempeños de los estudiantes, según veremos en el apartado que sigue. Este enfoque genera una apariencia de "objetividad", dado que se trabaja con fórmulas matemáticas, pero carece de significado sustantivo en términos de lo aprendido por el estudiante. Lo central para establecer una certificación no es qué peso relativo deben tener las actividades integradoras, las cotidianas y las pruebas, sino qué tipo de saberes y capacidades deben ser evaluadas a través de dichas actividades y constatadas en el desempeño de los estudiantes.

La "sacralización" de la nota numérica se puede observar también en la supuesta precisión de las mismas. Nótese, por ejemplo, que estas deben ser expresadas con hasta un decimal en El Salvador, Costa Rica, Chile y México y que varios reglamentos incluyen criterios para el "redondeo" o "aproximación" de los puntajes. La pregunta a hacerse al respecto es: ¿qué diferencia existe, en términos de dominio de la Matemática, entre un estudiante con una de nota de 6,8 y otro de 6,5? ¿Cuál es el significado sustantivo de cada una de estas cifras?

d. El "promedio" como juicio de valor sobre la persona del estudiante

Un cuarto elemento importante que emerge del análisis de los reglamentos es la orientación a establecer un "promedio" o calificación global para cada estudiante, en cada asignatura, cada dos o tres meses. El caso de México es especialmente ilustrativo: *"Promedio final de asignatura: Será el promedio de las calificaciones obtenidas en cada uno de los cinco bimestres que comprende el ciclo escolar"* (Art 10º). *"Promedio final de grado escolar: Será el resultado de sumar los promedios finales de las asignaturas y dividirlo entre el número total de asignaturas que se establecen para cada grado de la educación primaria y educación secundaria en el plan de estudios de educación básica"* (Art 11º).

La cuestión de los "promedios" tiene un doble efecto negativo -no nos referimos únicamente a los promedios matemáticos, sino a la práctica más amplia de "combinar" o "mezclar" diversos aspectos del desempeño de los estudiantes en una nota única-. En primer término, oscurece el significado de la calificación. Por ejemplo, en el caso de Uruguay se establece que *"La evaluación del proceso incluirá, además de la apropiación de los contenidos propios del curso, la apreciación de la actuación del alumno en relación a los siguientes aspectos: interés, actitud en el trabajo, integración social y asiduidad. Se tendrán en cuenta, además, los progresos educativos, la adquisición de valores y el desarrollo de aptitudes"* (Art. 27) y que *"El Profesor sintetizará la evaluación del proceso en un **juicio conceptual personalizado,** referido a los aspectos antedichos, y su **traducción numérica** en la **escala del 1 al 12**"* (Art. 29). La "síntesis" de tal

diversidad de aspectos en una calificación única impide que el estudiante y la familia sepa dónde está el problema de una calificación insuficiente: ¿en los valores, en el desarrollo de aptitudes, en la actitud en el trabajo o en la integración social? ¿Estoy siendo reprobado porque no me interesa la asignatura, porque tengo problemas para integrarme socialmente o porque no he adquirido ciertos valores? El afán de "integralidad" en la evaluación, mal implementado, puede dar lugar tanto a calificaciones sin significado concreto como a evaluaciones fuertemente disciplinadoras.

Algunos países han tomado nota de este problema y han tomado el recaudo de separar la evaluación de los comportamientos y de no incluirlos en la certificación, sino evaluarlos únicamente con fines formativos. Tal es el caso de El Salvador, donde se prevé la evaluación de aspectos de conducta, que forman parte de la asignatura Educación Moral y Cívica. Dichos aspectos son los siguientes:

1. Se respeta a sí mismo (a) y a los demás.
2. Convive de forma armónica y solidaria.
3. Toma decisiones responsablemente.
4. Cumple sus deberes y ejerce correctamente sus derechos.
5. Practica valores morales y cívicos.

Estos aspectos son calificados por separado. Para ello no se utiliza la escala de calificaciones numéricas, sino tres juicios de valor: Excelente, Muy Bueno y Bueno. Esta valoración, no es tenida en cuenta a los efectos de la promoción de los cursos.

Otra consecuencia de las calificaciones "promedio", "síntesis" o "juicios globales" es que terminan constituyéndose en una forma de "etiquetar" al estudiante. Recuérdese en este punto lo que planteamos en el capítulo anterior con respecto a las "atribuciones" que realiza el estudiante con relación a su desempeño. Siguiendo el planteamiento de Dylan Wiliam (2011), analizamos la diferencia entre atribución específica (*'tengo dificultades con la división de fracciones'*) y atribución global (*'no soy bueno para las matemáticas'*). El reporte a través de calificaciones globales opera en este último sentido: en lugar de enfocar al estudiante en lo que no está logrando y en lo que debe hacer para mejorar, se le coloca una etiqueta: *'soy un 1', 'soy insuficiente', 'soy bueno', 'soy 10'*. Las calificaciones globales, por tanto, no aportan al estudiante información específica sobre qué y cómo mejorar. En cambio, contribuyen a generar identidades estudiantiles negativas en aquellos que obtienen malas calificaciones.

Por otra parte, es inevitable preguntarse cuál es el sentido de asignar una nota global por trimestre y materia a cada estudiante, con una precisión de decimales. Desde nuestra perspectiva, se trata de una rémora del pasa-

do, en que los sistemas de calificaciones estaban orientados a promover la competencia entre los estudiantes y/o a procesos de selección para el acceso a estudios superiores. Solo en este último caso podría tener algún sentido establecer una calificación numérica global, siempre y cuando la medición fuese apropiada[7]. Tal como afirma Wiliam (2011), lo primero que miran los estudiantes cuando se les entrega un trabajo o un boletín son sus notas. Y lo segundo, inmediatamente después, son las notas de sus compañeros. Así como están definidas, más allá de las intenciones y de los discursos, las calificaciones globales solo sirven para promover la clasificación y comparación entre los estudiantes: los destacados, los buenos, los insuficientes, los mejores y los peores. Esto construye identidades estudiantiles, positivas y negativas. En definitiva, las calificaciones globales juegan el mismo papel que tanto criticamos en los rankings de países en las pruebas estandarizadas. Más adelante en este capítulo intentaremos formular algunas propuestas en torno a qué podríamos hacer los docentes al respecto.

2. ¿Cómo se construye una nota? La pócima de la bruja

En los trabajos de investigación que realizamos en varios países de la región tuvimos la oportunidad de dialogar con maestros y profesores acerca de los procedimientos y criterios que utilizan para calificar a sus estudiantes. La evidencia recogida, que presentaremos a través de testimonios tomados de las entrevistas realizadas con docentes[8], nos permitió arribar a dos conclusiones principales, que se pueden entender claramente a la luz de los reglamentos que acabamos de analizar. La primera está referida a los aspectos del desempeño de los estudiantes que son consideradas como relevantes para calificarlos. La segunda está relacionada con los procedimientos que se emplean para construir las calificaciones globales (bimestrales, semestrales o finales).

Con relación al primer aspecto, la constatación principal fue que se toman en consideración una enorme gama de aspectos, la mayor parte de ellos vinculados al disciplinamiento de los estudiantes más que al aprendizaje

7. Es técnicamente bastante complicado y sofisticado el proceso de construcción de escalas numéricas precisas para medir desempeños. Por otra parte, notas asignadas por profesores distintos sobre la base de criterios y procedimientos diversos, no son comparables entre sí. Sería necesario un proceso de estandarización.
8. Todos los testimonios incluidos en este capítulo están tomados de las entrevistas en profundidad con docentes de 6º de Primaria y 3º de Secundaria, realizadas en el marco de los estudios referidos en el capítulo 1.

propiamente dicho. A continuación explicamos este aspecto y lo ilustramos con testimonios aportados por los docentes entrevistados.

2.1. "Yo les califico hasta los respiros"

En general los docentes interpretamos el discurso normativo de la "evaluación integral" y de la "evaluación de procesos", en el sentido de que debemos calificar todo lo que hace el estudiante. Recuérdese el reglamento de Uruguay: *"La evaluación del proceso incluirá, además de la apropiación de los contenidos propios del curso, la apreciación de la actualización del alumno en relación a los siguientes aspectos: interés, actitud en el trabajo, integración social y asiduidad. Se tendrán en cuenta, además, los progresos educativos, la adquisición de valores y el desarrollo de aptitudes" (art. 27).*

En palabras de docentes de distintos países:

> *"Ellas conmigo tienen ocho, nueve o diez notas en un período, porque yo les califico hasta los respiros"* (entrevista C-21).

> *"Hago un promedio de todas las notas: disposición, memoria, redacción, habilidad para cortar, para pegar, evaluación escrita, participación, actitud, llegada temprano, comportamiento, trabajo en grupo. Para mí es un trabajo dispendioso porque hay alumnos que llegan a tener en un periodo 13 notas"* (entrevista C-17).

> *"Hago mucho énfasis en lo personal y social porque son factores importantes para el buen desarrollo de la clase. Si no ponen cuidado, disciplina y orden, tampoco* [lo harán] *en la toma de conocimiento"* (entrevista C-3) (esta docente asigna a estos aspectos un peso del 50% en la calificación).

> *"Todo es evaluable. Y la evaluación es constante [...] constantemente estoy observando. Todo no se anota, es bien difícil, lo ideal sería que se pudiese anotar todo, pero hay una ficha que me ayuda a anotar algunas cosas"* (entrevista P-9).

> *"Hay una ficha o lista de cotejo por alumno, dentro de los criterios para esta actitud ante el área y los valores. Por ejemplo 'levanta la mano para opinar' y tiene sus puntajes. Si el alumno recoge un papelito del suelo, se tiene en cuenta un tema transversal, la conciencia ambiental. Si participan en aula, en lluvia de ideas, también se les pone 'check'. Al final tienen un valor. 5 puntos: siempre participa; 3 puntos: a veces participa; 0: nunca participa"* (entrevista P-8).

> *"...todas [las evaluaciones] son para calificar... todas se califican y luego se promedian con la responsabilidad"* (entrevista P-15).

"En cualquier materia, por ejemplo, yo tomo un valor en el examen de siete puntos. Los otros tres puntos surgen de una escala que llevamos aquí, que considera muchos otros puntos, por ejemplo la asistencia, la disciplina, la puntualidad, higiene y varias actividades. Con eso les damos un porcentaje para completar los tres puntos que son, más los siete puntos del examen" (entrevista M-19) (recuérdese que la escala de calificaciones en México va de 5 a 10).

"Califico lo afectivo, la colaboración, la participación, el aseo, cómo se comportan, lo cognoscitivo" (entrevista G-14).

Como al final todo termina en una nota promedio bimestral, semestral o anual, por asignatura primero y, a veces, global para todas las materias (recuérdese el caso de México y véase más abajo la Figura 5.4 para Colombia), es muy difícil que alguien pueda saber a qué obedece la nota final. Como decíamos más arriba, una nota insuficiente puede significar que el estudiante no adquirió cierto conocimiento, que no levantó papelitos del suelo, que no vino a clase suficientemente aseado o que tiene que comportarse mejor. Con esto no pretendemos afirmar que no es importante valorar los aspectos actitudinales de los estudiantes, sino que no deberían formar parte de las calificaciones, y menos bajo la forma de un promedio con otras cosas muy diversas. Nos parece apropiado valorar lo actitudinal y reportar cada aspecto de forma clara por separado, sin consecuencias para la aprobación o reprobación del curso, al estilo de la propuesta de El Salvador que vimos más arriba.

Además de la ambigüedad del significado de las notas, y justamente como resultado de la mismas, estas se constituyen en un instrumento de disciplinamiento y motivación externa, que intenta sustituir el interés natural por aprender, como resultado de lo abstractas, descontextualizadas y poco significativas que suelen ser las actividades y tareas que proponemos en el aula.

2.2. "La nota es como la plata"

Las referencias al uso de la nota como instrumento de disciplinamiento y de motivación extrínseca aparecieron a lo largo de las entrevistas en todos los países. Pero la expresión más elocuente de esta práctica la aportó una docente chilena:

"Los alumnos solo toman en serio aquello que se califica... uno lo quiere hacer pero si esto no tiene notas, se bajan al tiro y hacen los trabajos a medias. Es como por plata, la nota es la plata. Y todo es para lograr algo, nada se trabaja porque sí" (entrevista CH-8).

Otro docente chileno explica cómo utiliza 'décimas' de puntos de calificación para motivar a los estudiantes: *"en todas las clases estaba preguntando para que se ganaran ellos sus décimas. Los chicos funcionan mucho con estímulo positivo de notas. Ojalá no fuera así pero en la casa si te portas bien tienes permiso para el cine y acá décimas de nota* [...] *tiene que haber una presión, algo escrito para que ellos se motiven a hacer algo..."* (entrevista CH-12).

En Colombia la mayoría de los entrevistados reconoció que, si bien el fin último de la evaluación es el aprendizaje, se suele emplear la evaluación -y el temor que esta produce en los estudiantes-, como herramienta para lograr "buena conducta", "responsabilidad" y "compromiso" con el estudio. Una docente lo expresaba de esta manera: *"eso hace que ellos trabajen y se afanen, porque saben que se van a ganar una buena nota"* (entrevista C-22). Otra docente expresaba: *"Cuando a un estudiante lo evalúan se compromete más. Hace un tiempo el gobierno decretó que los estudiantes no podían perder el año y ellas habían perdido el interés, entonces todo les daba lo mismo"* (entrevista C-14).

Tal como explicamos en el capítulo anterior citando a Shepard (2006), la nota se transforma en el objetivo de los estudiantes, desplazando a la motivación intrínseca por conocer más sobre el mundo y la cultura en la que viven. De todas formas, vale la pena reiterar aquí dos puntualizaciones realizadas en el capítulo anterior. En primer lugar, que la función de certificación es relevante en el sistema educativo, sobre todo a medida que se avanza en el mismo hacia los niveles superiores. En segundo lugar, que la certificación, además de una función social, cumple con una función de motivación extrínseca que es importante, en particular en el trabajo con adolescentes. El problema se genera cuando esta función se vuelve omnipresente, progresivamente hace desaparecer otras motivaciones y se transforma en el único objetivo de los estudiantes.

2.3. "El que hizo el trabajo de hormiga, ese pasa"

Otra forma de disciplinamiento es la vinculación de la nota al "esfuerzo" del estudiante. Las expresiones más elocuentes de este criterio de evaluación las encontramos en Uruguay, donde algunos docentes llegan a afirmar que el esfuerzo es más importante que el aprendizaje.

> *"Hay que hacer un seguimiento del estudiante del comienzo del año al final, no pretendiendo que todos lleguen al mismo punto, pero sí que haya habido una evolución desde marzo a noviembre"* (entrevista U-17). *"No me guío por el que le fue mejor, sino el cómo le fue a cada uno en relación a antes, yo creo que la mejor medida es uno mismo"* (entrevista U-12).

"Se valora el esfuerzo que hacen por tratar de llegar a un mínimo aceptable" (entrevista U-14). *"Tiene que haber tenido un trabajo continuo... El que hizo el trabajo hormiga, que se mantuvo todo el año, con altibajos pero todo el año trabajó, ese pasa"* (entrevista U-1). *"El grado de cumplimiento de tareas es evaluado, más allá de si está bien o mal la tarea, si intentó hacer algo, la información, si hizo una búsqueda"* (entrevista U-16).

Según es posible apreciar en estos testimonios, hay dos aspectos mezclados. En los primeros testimonios, el énfasis está puesto en el progreso de los estudiantes a lo largo del año con relación a su propio punto de partida. Este es un enfoque valioso de la evaluación, que analizaremos en el cierre de este capítulo. Requiere, como en la nueva experiencia que se está iniciando en Perú, establecer con claridad cuáles son los desempeños y/o logros relevantes y cuál es la progresión de niveles en cada uno de ellos.

El problema se plantea cuando la valoración del progreso se transforma en valoración del esfuerzo y, además, con independencia de que dicho esfuerzo conduzca o no a algún tipo de logro. Esto es lo que se puede apreciar en los testimonios ubicados en el segundo párrafo (U-1, U-14 y U-16).

También en Chile varios docentes expresaron tener en cuenta el esfuerzo a la hora de calificar, pero de un modo que aparece como más razonable. Utilizan la información sobre el esfuerzo para corroborar un mal resultado final. También para decidir si darle nuevas oportunidades a estudiantes que se han esforzado pero han tenido dificultades. Pero, nótese que no se trata de aprobar al alumno porque se esforzó, sino de habilitarlo a realizar un trabajo adicional o remedial.

"Tengo un tic [marca] de quien participó en clase o no. Si tengo alguien que nunca participó en clase ni hizo actividades y después tiene mala calificación, es una cosa de que no está poniendo su esfuerzo" (entrevista CH-6).

"Vienen y te dicen que tienen mala nota, si podrían hacer un trabajo o podría yo remediar. Conversas con ellos y transas, con algunos puedes hacer excepciones y con otros no, depende del desempeño que tengan en la asignatura y, sobre todo, del esfuerzo que ponen. Hay niños que les cuesta mucho cognitivamente pero se esfuerzan un montón, entonces tú a ellos también los ayudas" (entrevista CH-3).

Como suele ocurrir, hay un aspecto positivo y otro negativo en este enfoque. Por un lado, es sano inculcar valores vinculados con el trabajo y el esfuerzo para alcanzar ciertas metas. Por otro lado, no parece apropiado

valorar el esfuerzo por sí mismo si no conduce a logros concretos o si el esfuerzo se produce sobre temas y actividades que el estudiante no comprende. El resultado final es que algunos estudiantes, más que esforzarse en aprender algo a lo que no le encuentran sentido, lo que hacen es complacer al docente "haciendo como que" estudian y se esfuerzan. Los estudiantes que se rebelen contra propuestas de enseñanza irrelevantes correrán peor suerte. Por otra parte, tal como planteamos en el capítulo anterior, el esfuerzo, la concentración y la disciplina de trabajo son actitudes que normalmente emergen cuando uno está interesado en lo que está haciendo o aprendiendo.

3. Entre el promedio matemático y el juicio subjetivo

Decíamos al inicio del apartado anterior que un segundo hallazgo importante de los estudios que realizamos en varios países está vinculado con los procedimientos que los docentes utilizamos para construir las calificaciones bimestrales, semestrales o finales. Estos procedimientos varían entre dos posturas extremas y opuestas, que denominaremos como *"pseudo-objetiva matemática"* y *"subjetiva-arbitraria"*.

3.1. Los procedimientos matemáticos "pseudo-objetivos"

El primer tipo de procedimiento consiste en llegar a la calificación mediante operaciones matemáticas de distinto tipo, normalmente promediando una amplia diversidad de notas, en forma simple o ponderada.

Como vimos al analizar los reglamentos de evaluación, algunos países establecen ponderaciones matemáticas muy precisas: 30% para el trabajo cotidiano en el aula, 10% para las actividades realizadas fuera de clase; 50% para el resultados de al menos dos pruebas; 5% para el "concepto" del profesor y 5% para la asistencia. Este tipo de procedimientos generan una sensación de "objetividad matemática" al calificar y, en cierto modo, sirven para liberar al docente del estrés que generan tanto la decisión de aprobar o reprobar, como la duda sobre si está siendo justo con la calificación. Sin embargo, normalmente carecen de sentido sustantivo, porque no dicen nada acerca del tipo de saberes y desempeños que están en juego en cada uno de los componentes indicados. Veamos esto en un ejemplo concreto (Figura 5.3).

Figura 5.3: Estableciendo puntajes en una prueba de Lenguaje

IV PARTE. DESARROLLO. VALOR 18 PUNTOS

A continuación usted debe escribir un texto, aplicando cada uno de los objetivos temáticos escritos en la tabla. El tema es **"Amo la vida"**.

Tabla para calificar la redacción

Objetivos temáticos	3 puntos No errores	2 puntos 1 a 3 errores	1 punto 4 a 6 errores	Puntos obtenidos
1. Escritura legible en sus trazos. Repetición innecesaria de ideas.				
2. Uso del acento ortográfico: agudas, graves y esdrújulas.				
3. Uso correcto de la tilde diacrítica.				
4. Uso del hiato y ley del hiato.				
5. Uso correcto de las consonantes: b, v, c, s, z, m, n, h, rr, r, g, j. Uso gu con o sin diéresis.				
6. Utiliza signos de puntuación. Letra mayúscula.				

Fuente: Base de registros del estudio **La evaluación en primaria en ocho países de América Latina.**

No vamos a detenernos en el análisis de la parte sustantiva de la prueba, que exhibe una concepción bastante tradicional sobre la enseñanza del lenguaje, como estudio de las reglas gramaticales. Tampoco en la ambigüedad de la consigna -escribir un texto sobre el tema "Amo la vida"- que no ofrece al estudiante indicaciones sobre el tipo de texto esperado y su extensión.

Nos interesa que la atención del lector se enfoque en el modo de construcción de la calificación del trabajo. Según es posible apreciar, se trata de la parte IV de una prueba más amplia. A esta parte le corresponden 18 puntos en el total de la prueba. Los puntos se obtienen en función de la ausencia de errores gramaticales. Cualquier lector medianamente entendido en cuestiones de lenguaje, puede darse cuenta de que la probabilidad de cometer errores como los indicados es diferente para cada dimensión especificada. Por ejemplo, es más común que un niño tenga errores en el uso de la tilde diacrítica que en el uso del acento ortográfico. En segundo lugar, la probabilidad de ocurrencia de errores depende de la extensión del texto. En un texto breve siempre habrá menos errores. Si un estudiante escribe *"Amo la vida porque es hermosa"*, estaría cumpliendo con la consigna y debería recibir los 18 puntos, de acuerdo con la forma de calificar establecida, siempre y cuando escriba "amo" con mayúscula, "hermosa" con 'h', no ponga tildes y coloque

un punto al final de la frase. Otro estudiante, que escriba una carilla, seguramente tendrá un puntaje mucho más bajo porque cometerá algunos errores. En tercer lugar, el primer objetivo indicado, no condice con el resto. ¿Qué serían errores de "escritura legible"? ¿Y "repetición innecesaria" de ideas? De aquí surge otro consejo para nuestro estudiante '18 puntos': debe escribir en forma prolija (obviamente, con su frase "Amo la vida porque es hermosa", no corre el riesgo de repetir ideas).

Más allá del tono irónico que le hemos dado a este análisis y a lo extravagante de la propuesta, lo importante es que refleja un modo de proceder que muchos docentes seguimos en algunos momentos: establecemos un conjunto de actividades, preguntas o ejercicios, cada uno de ellos con un puntaje; corregimos los trabajos y establecemos el puntaje correspondiente a cada actividad; sumamos los puntajes para llegar a un total; finalmente, hacemos alguna suerte de equivalencia entre los puntajes y la escala de calificaciones. Normalmente lograr el 50% del puntaje equivale a una calificación de suficiencia. A veces llevamos ese porcentaje al 60%. Pero no realizamos una reflexión afinada con respecto a qué significa cada puntaje en términos de la relación entre el nivel de desempeño alcanzado por el estudiante y los propósitos educativos de nuestro curso.

Un ejemplo ilustrativo, a partir de la experiencia de elaboración de exámenes de Matemática en Uruguay, es el siguiente. Los exámenes son elaborados por los propios docentes en cada centro educativo. Es bastante frecuente que se incluya algún ejercicio especialmente original o difícil dirigido a los estudiantes destacados, para poder poner notas altas. El problema es que al hacerlo, se eleva la exigencia para llegar a la suficiencia (que suele ser resolver correctamente el 50% de la prueba) para todos los estudiantes, porque se sabe de antemano que la mayoría de los estudiantes no podrá lograr los puntos correspondientes a estos ejercicios especialmente complejos. En rigor lo que habría que hacer, si el objetivo central del examen es decidir si un estudiante está o no en condiciones de aprobar el curso, es ofrecer la posibilidad de realizar esos ejercicios más difíciles por fuera de la prueba principal, pero no considerarlos como parte del puntaje global de la misma.

Una vez asignada la calificación a una prueba o actividad extraescolar a través de un procedimiento de cálculo de puntajes, las cosas se complican más cuando se ingresa en el terreno de los promedios. En palabras de docentes chilenas, *"aquí una pasa al computador las notas que una lleva, el computador hace el promedio y una no hace nada"* (entrevista CH-8); *"en el software que trabajamos, ingresas todas las notas de todas las asignaturas y te entrega el informe por niño"* (entrevista CH-3). Dicho en otras palabras, las notas correspondientes a las diferentes actividades ingresan en una suerte

de "licuadora" que las promedia, dando como resultado un número mágico final de significado incierto.

En Colombia, a pesar de la "descentralización" de los reglamentos de evaluación, la cuestión de los promedios parece estar fuertemente establecida[9], con ponderaciones específicas para la amplia diversidad de aspectos que son calificados y que describimos un poco más arriba.

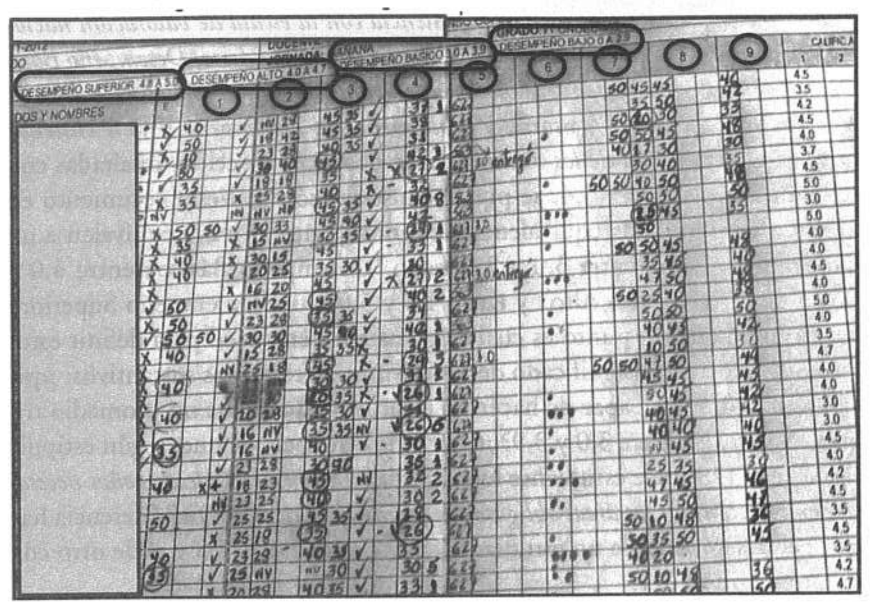
> *"Lo que se llama valoración académica, que es la parte de conocimientos en sí, tiene un valor del 30%. La valoración actitudinal vale también el 30%, el examen parcial vale 20% y el examen final vale 20%. Son muchas cosas que se le evalúa al niño, ellos no pueden decir 'ah yo perdí porque nunca me recogió el cuaderno', uno tiene muchas maneras de evaluar a estos niños"* (entrevista C-22).
>
> *"Por período hay una evaluación parcial y una final, eso tiene un valor del 50% de la nota final. El otro 50% está repartido, un 65% que es la parte cognitiva, hablamos de los trabajos, actividades, talleres, todo lo que se trabaja en la clase, laboratorios; y el 35% es la parte afectivo-espiritual que maneja el colegio [...] son los porcentajes que maneja la institución y eso no se puede cambiar"* (entrevista C-4).

Aparentemente, la diversidad de propuestas entre colegios está vinculada principalmente a los pesos relativos en la nota final de los distintos aspectos considerados, pero la lógica central de construir la calificación a través de promedios matemáticos es más o menos la misma.

Con el fin de ilustrar a qué tipo de prácticas conduce esta lógica, en la Figura 5.4 se incluye una planilla de registro de calificaciones de Colombia (se eliminaron los datos identificatorios de la institución).

9. De todas formas, cabe reiterar que los estudios realizados no tienen carácter representativo y no son generalizables.

Figura 5.4: Planilla de registro de calificaciones en Colombia



Fuente: Base de registros del estudio **La evaluación en primaria en ocho países de América Latina.**

Las columnas numeradas del 1 al 9 en la parte superior (indicadas en círculos), corresponden a las siguientes dimensiones:

1. evaluación escrita;
2. participación en clase;
3. disposición actitudinal;
4. asistencia;
5. sustentación salida al tablero (pizarrón);
6. disciplina;
7. trabajo en equipo;
8. coevaluación;
9. trabajo en clase.

Las columnas que se ven luego hacia la derecha (solamente la 1 tiene datos) corresponden a los promedios trimestrales. Lo primero que uno se pregunta, a la vista de esta planilla, es qué significado puede tener cada una de las notas colocadas en dicha columna: 4.5, 3.5, 4.2, 4.5, 4.0, 3.7. Más llamativa aun resulta la equivalencia que se establece entre la escala de calificaciones, que va de 0 a 5, con la escala oficial nacional. Recuérdese que el artículo 5º del reglamento del Ministerio de Educación (Decreto 1290) establece que *"Cada establecimiento educativo definirá y adoptará su escala de*

valoración de los desempeños de los estudiantes. Para facilitar la movilidad de los estudiantes, se deberá expresar la equivalencia con la escala de valoración nacional: Desempeño Superior; Desempeño Alto; Desempeño Básico; Desempeño Bajo. Desempeño básico se entiende como la superación de los desempeños necesarios en relación con las áreas obligatorias y fundamentales. Desempeño bajo se entiende como la no superación de los mismos". Si se observan las celdas resaltadas con recuadros en la Figura 5.4, se puede constatar cómo el establecimiento en cuestión definió dicha equivalencia. Las notas entre 0 y 2,9 equivalen a un 'Desempeño Bajo'; entre 3,0 y 3,9, a un 'Desempeño Básico'; entre 4,0 y 4,7, a un 'Desempeño Alto'; y entre 4,8 y 5,0 a un 'Desempeño Superior'. La pregunta en este punto es cuál es el sustento empírico para definir estos puntos de corte y el significado de los mismos en términos sustantivos: ¿qué cosas aprendió y es capaz de hacer un estudiante que tiene un promedio trimestral ubicado entre 3,0 y 3,9? ¿Cuál es la evidencia de que, según estipula el Decreto 1290, ese estudiante logró *"la superación de los desempeños necesarios en relación con las áreas obligatorias y fundamentales"*. ¿Qué diferencia hay entre el desempeño de un estudiante cuyo promedio es 2.8 y el de otro con un promedio de 3.1?

3.2. La calificación como clasificación "subjetiva-arbitraria" de estudiantes

Muchos docentes tienen conciencia de la ausencia de sentido de este tipo de promedios matemáticos, por lo que prefieren seguir un camino más "conceptual". El problema es que, por lo general, esto se traduce en una suerte de clasificación subjetiva y poco transparente de los estudiantes. Como, según vimos en el capítulo anterior, los docentes tendemos a pensar nuestros cursos en términos de "temas a dar" y nos cuesta pensar en términos de qué esperamos que los estudiantes aprendan, la "evaluación conceptual" suele transformarse en una suerte de clasificación subjetiva no explícita de los estudiantes, cuyo significado es difícil de comunicar. Los testimonios que siguen ilustran este enfoque.

> *"Tenés en cuenta todo para la calificación final, pero no es un promedio. Uno no le pone al esfuerzo una nota, hay cosas que no podés hacer un promedio porque no son un número. Esfuerzo es un concepto del estudiante, más allá de los contenidos. Incluye todo lo que es el estudiante, no es un número, no podés hacer promedio con cosas que son subjetivas. Las tenés que evaluar y no entra en el promedio matemático"* (entrevista U-2).

"Para calificar uso las evaluaciones de todos los días. La nota que va al boletín es el resumen de todo lo que yo sé del alumno. No hago nada en forma especial" (entrevista U-15).

"Lo hago subjetivamente porque además depende para quién: un 6 para vos no es lo mismo que un 6 para mí. Hay un grado de aceptabilidad en función de características de cada uno: no puedo hacer pesar más el escrito a alguien que tiene problemas para expresarse por escrito, sino que le hago pesar más lo oral de todos los días. Es más, lo hago al revés, no sé si esto es correcto o no, tengo la visión global de 'yo pienso que le tengo que poner tanto', y cambio la nota matemática en el oral para que después le dé eso y no me joroben con lo matemático. Si le va bien en el escrito y no hace nada, le amarreteo en el oral para que haga, su aceptabilidad está en un nivel diferente" (entrevista U-9).

"El sistema me exige evaluaciones y notas, aunque yo soy muy de que a veces evalúo al ser humano sin tener en cuenta las planillas, sé quién pasa y quién no, a quien dejamos un tiempo o ponemos a que se esfuerce más" (entrevista C-11).

"A veces lo hago y a veces no lo hago. Yo no necesito llevar una planilla con tantos números... yo te puedo decir a quien le puedo poner un 5 y a quién le puedo poner un 2 sin que ellos me respondan" (entrevista C-17).

Si se analizan estos testimonios con cuidado se puede apreciar que, en el fondo, este modo de definir las calificaciones adolece de problemas similares a los promedios matemáticos: el significado de las mismas sigue siendo opaco para los estudiantes e implica también la combinación, en un solo juicio o concepto, de una cantidad heterogénea de dimensiones.

Por otra parte, expresiones como *"la nota es el resumen de todo lo que yo sé del alumno"; "yo pienso que le tengo que poner tanto"; "yo sé quién pasa y quién no"* o *"yo te puedo decir a quién le puedo poner un 5 y a quién le puedo poner un 2 sin que ellos me respondan"*, reflejan una visión de la función de calificación que, por lo menos, es discutible. Muchos docentes, seguramente sin ser del todo conscientes, etiquetamos a nuestros estudiantes rápidamente. ¿Con qué frecuencia pensamos en términos de 'este es un 5', 'este es un 8', 'este es un 10'? Si muchos docentes "ya saben" a mitad del año quiénes van a aprobar y quiénes no, ¿no hay nada que pueda hacer el estudiante? ¿Cuál es entonces el sentido de enseñar y aprender?.

Esta manera de abordar las calificaciones refuerza su carácter de instrumento de disciplinamiento, en el mal sentido del término. Tal como afirmaba Wiggins (1998) en su paralelismo entre las evaluaciones y el juego de *"frío, tibio, caliente"*, al estudiante no le queda otro camino que intentar

adivinar qué quiere su profesor, e intentar complacerlo, intento que muchas veces está condenado al fracaso desde el comienzo, porque el profesor "ya sabe" que ese estudiante 'es un 1'.

En este punto es importante señalar que el problema no es la subjetividad, sino la arbitrariedad. Muchas veces tendemos a confundir ambos términos. El Diccionario de la RAE define como *"arbitrario"* a aquello que está *"sujeto a la libre voluntad o al capricho antes que a la ley o a la razón"*. Según se puede apreciar, el foco de la definición está puesto en el "capricho" y en colocarse por fuera de alguna forma de ley o razón. En el caso de la evaluación, la valoración del docente es arbitraria cuando no está basada en criterios explícitos y conocidos por los estudiantes. En toda evaluación existe siempre algún grado de apreciación subjetiva y/o inferencia valorativa, tal como fue expuesto en el capítulo anterior. No hay forma de evitarlo[10]. Pero se puede evitar la arbitrariedad, sin dejar de lado la subjetividad. El modo de hacerlo es explicitar claramente los criterios de valoración que empleará el docente. Para ello es necesario utilizar alguna herramienta, como las "rúbricas", que explicite las intenciones educativas, sus diferentes niveles de logro y el modo en que el docente construye su valoración del desempeño del estudiante. Sobre estos temas, en especial sobre el uso de rúbricas con fines de calificación, volveremos más adelante en este capítulo.

3.3. A modo de resumen

Los reglamentos de evaluación en cada país establecen la escala que deben utilizar los docentes para la evaluación con fines de certificación, es decir, para calificar a los estudiantes. Normalmente establecen también una "nota de suficiencia" cuyo significado específico queda librado a la interpretación de cada docente y/o a las definiciones curriculares incluidas en los programas de estudio. Como estos últimos no suelen definir niveles de logro vinculados a la escala de calificaciones, se genera un espacio de ambigüedad. Algunos reglamentos intentan cubrirlo apelando a la definición de procedimientos y ponderaciones matemáticas que deben ser seguidas para establecer una calificación.

Dentro de este marco normativo, los docentes nos movemos entre dos posturas extremas a la hora de definir las notas de los alumnos. Algunos lo hacemos asignando puntajes numéricos a las distintas actividades de una

---

10. Incluso cuando se procede a través de promedios y cálculos matemáticos, existe cierto grado de subjetividad -y a veces de arbitrariedad- en el momento de definir el puntaje que corresponderá a cada actividad o el contenido mismo de la actividad.

prueba, otorgando notas a distintas actitudes, comportamientos o acciones de los estudiantes, para luego hacer un promedio más o menos exacto (el grado de apego al resultado matemático preciso depende de cada docente) de los "números" obtenidos de esa manera. Como tales promedios suelen incluir aspectos muy diversos -esfuerzo, participación en clase, comportamiento, resultado de pruebas, tareas realizadas fuera del aula, entre otras- su significado final es incierto.

Otros docentes optamos por evitar la asignación de puntajes y la realización de promedios, y calificamos a partir de la opinión que nos vamos formando de cada estudiante a lo largo del curso. En el fondo, se trata de una suerte de "promedio subjetivo" que cada uno hace interiormente, en el que los aspectos considerados y el peso relativo de cada uno no son explícitos. De allí que la nota resultante pueda ser calificada como arbitraria.

En ambos casos, el resultado final es la expresión de un juicio de valor global sobre el estudiante, a través de una calificación que incluye aspectos disímiles que deberían ser reportados por separado, y que no ofrece ni al estudiante ni a las familias información sobre qué es necesario mejorar y cómo hacerlo.

Esta situación es particularmente compleja en la educación media, en la que los estudiantes son evaluados por una gran cantidad de docentes, cada uno con una visión diferente acerca de qué aspectos valorar y cómo "ponderarlos". Como resultado de ello, una de los desafíos principales del "oficio de estudiante" es aprender a decodificar lo que cada docente espera y actuar en consecuencia. Algunos estudiantes lo hacen muy bien, se adaptan y sobreviven. Pero otros no logran terminar de entender qué quieren ciertos docentes ni qué es lo que ocurre a su alrededor con las evaluaciones, y terminan fracasando en sus estudios.

El efecto negativo más preocupante es que, en este marco de ambigüedad, las calificaciones se separan de los aprendizajes y pasan a tener "valor propio", se constituyen en un fin en sí mismas y en el objetivo de los estudiantes: estudian para alcanzar una determinada calificación y actúan para conformar al docente. Las calificaciones se transforman así en un instrumento de disciplinamiento, desalientan a los alumnos de menor rendimiento y son generadoras de fracaso.

En términos de Wiggins, *"el problema que necesitamos solucionar no es si usamos letras, números o juicios para calificar.... Los símbolos no son el problema. El problema es la falta de puntos de referencia estables y claros al utilizarlos... El problema es el hábito de usar una sola calificación, que no tiene un significado claro, acordado y estable, para resumir todos los aspectos de un desempeño complejo... Necesitamos calificaciones mejor definidas, y necesitamos más tipos de ca-*

lificaciones, no menos... Las letras o los números, como todos los símbolos, ofrecen un modo eficiente de resumir información compleja... es tarea del docente darle sentido presentando hechos, juicios, diagnósticos y orientaciones en un formato amigable y documentado" (1998: 247). Por otra parte, la práctica de resumir desempeños complejos en un único juicio de valor global, se transforma en una valoración de la persona del estudiante más que en información sobre su desempeño en torno a actividades y saberes específicos.

Para cerrar esta parte con algo de humor e ironía, ya en 1957 Paul Dressel definía a las calificaciones de la siguiente manera: *"Una calificación puede ser considerada como un reporte inadecuado de un juicio impreciso, realizado por un juez sesgado y variable, acerca del grado en que un estudiante ha logrado un nivel indefinido de dominio, de una proporción desconocida de un material indefinido"* (citado por Wiliam 2011, 123).

4. Calificaciones y modelos educativos

Superar los problemas involucrados en las prácticas de calificación vigente es sumamente complejo, porque las mismas forman parte central de un modelo educativo creado en el marco de la revolución industrial que, si bien ha tenido algunas transformaciones menores, mantiene su esencia. Las calificaciones forman parte del ADN de este modelo que, en definitiva, es un sistema cultural[11]. De allí las dificultades para modificarlas. Para dar cuenta de la complejidad del problema, hemos organizado este apartado en dos partes. En la primera analizamos las características del modelo escolar del que las calificaciones, tal como las conocemos y practicamos, forman parte. En la segunda presentamos un modelo escolar alternativo, denominado "educación proyectiva", que fue desarrollado en Uruguay por un grupo de maestros (Ferrer, L. 2006). Al interior de este modelo, la evaluación se organiza de un modo radicalmente diferente. En el apartado siguiente, teniendo en cuenta que por el momento los docentes debemos continuar trabajando dentro del modelo existente, formularemos algunas pistas de trabajo para mejorar nuestra manera de calificar en el marco del sistema actual (mientras esperamos que se produzcan transformaciones más profundas en los sistemas educativos...).

11. Recuerde el lector el análisis sobre la enseñanza como práctica "cultural" desarrollado en el capítulo 1.

4.1. Las calificaciones y el modelo escolar de la sociedad industrial

La manera en que calificamos a nuestros estudiantes es una rémora de los orígenes de los sistemas de educación formal, desarrollados en la segunda mitad del siglo XIX y primera del XX, en que adquirieron sus características fundamentales, derivadas de las necesidades del modelo de desarrollo industrial (y de las nacientes democracias): la organización de los estudiantes en grupos de clase fijos por edades, ubicados espacialmente en un salón para cada grupo, en los que todos realizan las mismas actividades en los mismos tiempos, con contenidos organizados en programas por disciplinas o "materias" que deben ser cursadas y aprobadas en forma secuencial y compartimentada. Los estudiantes son evaluados en forma frecuente, con el fin de asignarles una calificación a partir de la constatación del dominio de los contenidos estipulados. En función de ello, al final de cada año se decide qué estudiantes están en condiciones de avanzar al grado siguiente. A todo esto se agrega, según mostramos en el capítulo 2, el carácter básicamente memorístico y declarativo de las evaluaciones y de la relación con el conocimiento que las mismas promueven. Leonardo Ferrer en su libro Educación Proyectiva (2006), analiza con agudeza el vínculo entre el modelo escolar y las necesidades de la sociedad industrial.

> *"La educación formal elemental tuvo pues por principales finalidades imponer los valores sociales y personales así como los consecuentes hábitos acordes a las necesidades estructurales del colectivo industrial, e impartir masivamente la instrucción básica necesaria para que ese sector amplio de la población funcionara como base de una comunidad de producción en serie de bienes y servicios... Desde la escuela universal se propagaron valores como la disciplina, la obediencia y principalmente la conformidad con la condición de ser proyectado, solo limitada por la conciencia del derecho a recibir lo necesario para la subsistencia... Por supuesto entre unas escuelas y otras hubo y hay diferencias: mientras en unas las aulas a las que se constriñe a niños y jóvenes son espaciosos salones donde se respira aire acondicionado, en otras entra frío por los vidrios faltantes. Mientras en unas a determinadas conductas le corresponde un llamado de atención, en otras un fuerte castigo, y mientras en unas se saluda poniéndose de pie a la "Señora Maestra" en otras se la tutea. Es más, mientras en unas los mensajes se memorizan en otras se razonan. Pero la escuela industrial fue y es esencialmente una, aun si se presenta bajo diferentes formas según los diferentes contextos...* (Ferrer, 2006: 44-45).

> *"El ambiente educativo de la escuela habitual se condice con su finalidad. La relación entre alumnos y docentes es siempre vertical, el docente es dueño*

de un programa. Las relaciones personales entre alumnos son prácticamente inexistentes, limitándose a las que se producen durante trabajos colectivos predeterminados en su forma y circunstancia, o a los tiempos "de recreo". En el curso de las actividades programadas los alumnos están compartimentados en grupos por edades y recluidos en aulas que rodean a un patio acondicionadas para las relaciones de aprendizaje mencionadas. En esta circunstancia los educadores imparten a los alumnos -constituidos en sujeto colectivo- un programa que se entiende contiene aquellos objetos de aprendizaje que deben necesariamente ser abordados por todos los educandos, en una secuencia y con una profundidad relativamente predeterminadas. Aún reconociendo que el programa pueda desarrollarse atendiendo a necesidades o deseos particulares de los alumnos, en esta instancia el educador juega un rol determinante. Es quien propone (impone) la temática a tratar, quien determina la profundidad y el momento en que la misma será tratada, así como el carácter con que se abordará. El protagonismo del educando está limitado por la propia constitución de la dupla educando-educador" (Ferrer, 2006: 46-47).

La calificación como "juicio" que resume una valoración/etiqueta de cada estudiante es funcional a este modelo de escuela, al igual que la definición de criterios de promoción de grado establecidos en función de la aprobación individual de cursos compartimentados. La calificación funciona, a la vez, como un instrumento disciplinador y como un procedimiento de clasificación y ordenamiento de los estudiantes. Este último elemento, el relativo al ordenamiento, muchas veces es olvidado. En el origen de la amplitud de las escalas de calificaciones y de su supuesta precisión hasta los decimales estuvo, a lo largo de buena parte del siglo XX, la intención de ordenar a los estudiantes para procesos de selección para estudios superiores, así como de promover la competencia entre ellos por destacarse con las "mejores notas". Esto generó entre los docentes una cultura de la precisión en el ordenamiento de los estudiantes a través de las calificaciones que, según mostramos en este capítulo, se ha transmitido de generación en generación y perdura, en forma solapada, hasta hoy.

Para comprender mejor el problema nos detendremos un momento a mostrar cómo plantea Ferrer su visión de un modelo de escuela alternativo, que permita asumir los desafíos derivados de las profundas transformaciones en la sociedad contemporánea. El eje central de un nuevo modelo debería ser el desarrollo de la capacidad de proyecto de las personas. El autor reinvindica conceptos impulsados por movimientos de cambio como la Escuela Nueva: la enseñanza personalizada, el interés como motor del aprendizaje, el carácter social del proceso de aprendizaje, el desarrollo del espíritu crítico y creativo,

el valor de lo lúdico en el aprendizaje, entre otros. Sin embargo, enfatiza que el fracaso de la mayoría de los intentos de transformación educativa, radica en que no se han modificado las estructuras básicas de la organización de la actividad educativa señaladas más arriba. *"Quienes tenemos años de docencia sabemos lo difícil o más bien imposible que es sobrepasar lo discursivo y hacer realidad concreta dichas aspiraciones y puntos de vista en los marcos de la educación habitual. Esto no se debe a la falta de voluntad ni de creatividad de los educadores prácticos;* [el problema central] *radica en la falta de un marco pedagógico y orgánico adecuado a la materialización de las aspiraciones antedichas"* (Ferrer, 2006: 16-17).

Para fundamentar la necesidad de un modelo escolar alternativo, Ferrer describe en los siguientes términos los grandes cambios sociales en curso:

> *"A partir de la segunda mitad del siglo XX acaeció una revolución científico tecnológica tanto o más importante que la industrial del siglo XIX. Surgieron robots capaces de producir objetos en serie a un costo mucho menor (es decir disminuyendo trabajo humano) que el de las tradicionales máquinas industriales... Con el desarrollo de estos nuevos medios productivos se hicieron posibles unidades descentralizadas, integradas a los sistemas tradicionales de producción, y el desarrollo paralelo de producciones materiales y de servicios a escala relativamente menor, adaptados a necesidades y gustos diversos, capaces de responder a necesidades particularizadas, específicas, para las que la producción industrial tradicional no tiene respuesta... Acaeció también la extinción de infinidad de trabajos... no solo se simplificaron procesos productivos, sino que una serie de problemas... pudieron ser eliminados como tales en un laboratorio mediante equipos de investigación intelectual. La informática eliminó también trabajo, ahora ya no físico... sino mental... Las nuevas formas de producir condicionan nuevas formas de relacionamiento... Cada vez más la producción tiende a realizarse a distancia y las personas pueden desarrollar su vida productiva en los suburbios. Tiende también a desaparecer la división tajante entre «intelectuales» y «trabajadores»; manos y cerebros electrónicos liberan del trabajo rutinario a un creciente número de personas, reduciendo cada vez más la función humana a la investigación y la toma de decisiones, a la gestión...*
>
> *Con el advenimiento de los equipos informáticos (medios de producción material y pensante neo-artesanales, desconocidos en las sociedades industriales tradicionales cuyo alcance revolucionario en términos sociales recién comienza a vislumbrarse) comenzaron a surgir nuevas ocupaciones; muchas personas comenzaron a organizarse en pequeñas empresas oferentes de servicios y productos finales, o de servicios a la industria, a los aparatos burocráticos, e incluso al propio sistema educativo. La producción en serie*

hasta donde podemos imaginárnoslo continuará existiendo pero resolviendo una parte en términos relativos cada vez menor de los productos finales (de consumo)... Con el desarrollo de la colectividad gestionista tienden a extinguirse los gustos y los hábitos masivos. La posibilidad de acceso a bienes personalizados desarrolla la diversidad de gustos, de necesidades espirituales... La globalización de la cultura que conecta a la persona con diferentes hábitos, valores, etc., la generalización de una cultura relativamente superior y, particularmente, la atomización de los centros de poder, desestiman el dogma, abriendo paso a la conciencia crítica. Con el post-industrialismo se produjo una reacción antidogmática..." (Ferrer, 2006: 48-51).

A estos elementos podrían agregarse otros que han tomado fuerza a lo largo de la última década (recuérdese que el libro de Ferrer fue escrito en el año 2006, cuando los celulares inteligentes apenas estaban emergiendo) y que operan en la misma dirección: el desarrollo de las redes sociales como medio de comunicación y de participación social y política, así como la explosión de las llamadas aplicaciones y los trabajos vinculados a ellas, tanto en la programación de las mismas (desarrolladores) como en su instrumentación (servicios de transporte, alojamiento, turismo, distribución de comidas, espectáculos, libros electrónicos, podcasts, radios y canales de televisión por internet, cursos y conferencias virtuales de todo tipo, por mencionar algunos).

4.2. La educación proyectiva, un modelo alternativo

En este nuevo contexto, sostiene Ferrer, la educación formal debería re-formularse en torno a una idea-eje central: el desarrollo de la capacidad de proyecto de niños y jóvenes. En este sentido, comienza por caracterizar al ser humano como un "ser proyectivo": *"El hombre se distingue de las demás especies por su capacidad de realizar proyectos, de imaginar realidades y construirlas; pero esta capacidad no se reproduce automáticamente como la de cazar de los gatos, o de construir nidos de los pájaros. Se realiza mediante la educación...* (Ferrer, 2006: 14).

La capacidad de proyecto implica trabajar en equipo, planificar, co-gestionar, proponer, decidir. *"La sociedad gestionista a la que nos asomamos valora más la autodisciplina que la obediencia, el compromiso que el cumplimiento del deber. Sin compromiso, sin autodisciplina, sin espíritu creativo, sin empatía, no hay proyecto. Educar para el proyecto no supondrá pues «rescatar» sino consolidar una nueva escala de valores que emerge de las relaciones sociales de co-gestión, de con-vivencia"* (Ferrer, 2006: 53). El autor destaca los siguientes valores como propios de la educación proyectiva:

- *El espíritu creativo -que incluye al crítico- es para un colectivo gestionista uno de los valores fundamentales, una actitud que estimamos central. Entendida pues como un elemento de la proyectividad, la creatividad que propugnamos se «desmarca» de la idea de actividad irreflexiva, intuitiva, superficialmente espontánea y solo propia de espíritus «artísticos» que muchas veces se le endilga. En efecto, el proyecto supone lo creativo en cualquier área, desde la de las matemáticas hasta la de los negocios, el deporte o las propias artes. En todos los casos, la acción creativa, aun cuando parte de inferencias o imágenes más o menos espontáneas, se apoya en la información, la reflexión, el trabajo sistemático* (53).
- *La idea de disciplina ligada tradicionalmente a la de disciplinamiento se opone a la de libertad y creatividad. Sin embargo la disciplina entendida como autodisciplina es condición necesaria de la acción creativa. Libertad y disciplina, lejos de ser antinómicas, son dos valores que se retroalimentan... La falsa antinomia tiene origen en la idea de disciplina como imposición. Quizás desaparecería si en lugar de disciplina dijéramos autodisciplina. Entendida como autodisciplina ella no es atentatoria al desarrollo de la libertad; todo lo contrario, es una herramienta que permite el desarrollo del proyecto como realización de libertad* (54).
- *Un valor fundamental en una sociedad gestionista es el del compromiso. Los modos de producir que vienen ocupando y ocuparán en no mucho tiempo a la absoluta mayoría de las personas, requieren ya no de mera responsabilidad sino de compromiso. Como vimos, las actuales unidades gestionarias tienen una gran movilidad interna, condicionada por la rápida evolución de conocimientos y recursos disponibles, así como por la permanente evolución de las demandas que emergen de un contexto que hoy es mundial. Este modo de interrelación exige de la persona elasticidad, versatilidad y del permanente enriquecimiento de los conocimientos, del ajuste continuo de su aporte, del permanente acomodamiento a transformaciones organizativas. Para ello es imprescindible el compromiso* (54).
- *Para la actividad de gestión la inteligencia emocional es tan importante como la intelectual. En el curso de desarrollo de un proyecto la carga emocional, tanto en lo referente al entusiasmo como a la capacidad de superación de la frustración o de control de la ansiedad ante lo contingente, es un factor tan determinante como la claridad de objetivos y el dominio de medios. En tanto participar de proyectos requiere de la capacidad de co-elaborar, de relacionarse con mentalidad de asociado, la cultura proyectiva supone el desarrollo de la inteligencia social, es decir*

de la capacidad de entenderse y entender al -otro (de ponerse en el lugar del otro y así co-mentar). Esta es una condición imprescindible para la co-gestión... (55-56).

Paralelamente, el autor plantea una nueva relación con los saberes o contenidos curriculares en el marco de la Educación Proyectiva. *"En la actividad gestionista se impone la permanente adquisición de nuevos conocimientos, pero ya no solo de los tradicionales conocimientos especializados sino de un conocimiento substancial y transversal. En lo que se refiere al pensamiento, se requiere un modo de pensar operativo, explicativo, dialéctico, estructural, en lugar del pensamiento lineal y estático que requería la cultura de «masas»"* (55). Esto implica combinar dos tipos principales de espacios formativos en la escuela proyectiva: los "*programáticos*", orientados a la transmisión del saber social acumulado, y los "*facilitados*", orientados a trabajar en torno a proyectos, a partir de la iniciativa de los estudiantes, en los que los conocimientos a trabajar dependen de los requerimientos de dichos proyectos.

En las **instancias "programáticas"** el objetivo fundamental es propiciar el ingreso de niños y jóvenes *"en un universo de ideas y conocimientos que trascendiendo al de su entorno cultural y a su interés inmediato, les dé un marco adecuado a experiencias y aprendizajes autónomos... La educación formal no puede dejar la cuestión del interés librada a la suerte, limitándose a facilitar demandas. Tampoco es posible que en la práctica espontánea se presenten en orden las circunstancias de aprendizaje que vayan habilitando a nuevas adquisiciones de conocimiento así como a los sucesivos requerimientos de la vida social... La institución educativa debe incluir pues instancias programáticas, es decir una secuencia «artificial» de aprendizajes. De otro modo sería imposible la transmisión del conocimiento social acumulado. En la instancia de aprendizaje programático el objeto, el ritmo y el modo de aprendizaje es propuesto por la institución. Ella fija entre otros aspectos las reglas del proceso, la integración y funcionamiento del grupo"* (73). Obviamente, dado que a los aprendizajes programáticos se les destina solamente una parte de la jornada escolar, se hace necesario ser especialmente selectivo. Es necesario enfocarse en las materias y temáticas que ofrecen mayores posibilidades de involucrar y motivar a los alumnos y que, al mismo tiempo, ofrecen más posibilidades de desarrollar las facultades proyectivas.

En paralelo con las progamáticas, las **instancias "de facilitación"** tienen como objetivo que los propios estudiantes sean los que definen el contenido de las actividades, a partir de la generación de proyectos. *"A proyectar se aprende proyectando. En esta instancia la intervención del educador debe ser la de facilitar las actividades proyectivas propuestas por el educando. En el aprendizaje facilitado el objeto, el ritmo y - en cierta forma- el modo de aprendizaje es*

que están realizando en las jornadas escolares, sus necesidades de conocimiento para avanzar en sus proyectos. *"El perfil proyectivo es necesariamente un perfil reflexivo. Tanto el aprendizaje programático como el facilitado deben complementarse con instancias de reflexión personal a las que la institución educativa debe asignarles espacio, tiempo y facilitación"* (76).

Según surge de la enumeración anterior, la organización de tiempos, espacios y agrupamientos de los estudiantes en este nuevo modelo escolar es radicalmente diferente al vigente en las escuelas a las que estamos habituados. *"La institución proyectiva debe tener una estructura física y temporal tal, que el niño o el joven tengan un dónde y un cuándo para sus actividades proyectivas"*. Este es un elemento clave de la propuesta: nuevas finalidades y propósitos educativos requieren de nuevos formatos escolares. Según se afirmó más arriba, los intentos de incrustar nuevos modos de enseñar y otros cambios educativos en una estructura concebida para otros fines, está en la base de buena parte de los fracasos de dichos cambios. Esto incluye, por cierto, a las calificaciones, según veremos enseguida.

La organización de tiempos, espacios y agrupamientos en la escuela proyectiva tiene tres características distintivas principales (Ferrer, 2006):

a. *"Un modelo de reformulación educativa incluye en primer lugar la propia arquitectura escolar (o en su defecto una utilización alternativa de los espacios de la escuela habitual)... cada una de las vivencias que hacen a la realización del niño o del joven como ser proyectivo, las de reflexión personal, las de encuentro social, las de creación artística, las de investigación, deben tener espacios propios. La institución educativa debe ofrecer no solo espacios infraestructurados para actividades específicas, sino la posibilidad de inventar espacios individuales o colectivos"* (76).
b. *"El ambiente pedagógico debe dar lugar a la mayor cantidad de variantes de relacionamientos, tanto en la forma como en el contenido. Relaciones en las que se compartan objetivos, en las que se ayude y se sea ayudado, en las que se enseñe y en las que se aprenda"* (66). El autor enfatiza especialmente la necesidad de interactuar en distintos agrupamientos de estudiantes y con la diversidad de adultos educadores de la institución, en distintos momentos, por oposición a la práctica habitual de pertenecer a un único grupo de estudiantes y de interactuar en forma casi exclusiva con un maestro o con los profesores de las disciplinas, y siempre en el ámbito del aula.
c. *"La institución debe ofrecer la posibilidad de gestionar tiempos propios. Fuera de aquellos en los que se desarrollan actividades programáticas,*

propuesto por el alumno; el sujeto de aprendizaje, en caso de ser colectivo, es auto generado [esto implica que los estudiantes conforman por su propia iniciativa grupos diversos de los habituales, con diferentes edades, en torno a un interés compartido]. *El objetivo de esta modalidad de aprendizaje es que el alumno realice experiencias directas de tipo proyectivo en relación a diversos objetos culturales... En la instancia de aprendizaje facilitado el alumno hace conciencia de la relación entre fines y medios, aprende a considerar los instrumentos materiales y espirituales requeridos para concretar un objetivo, fortalece la capacidad volitiva, aprende a preguntar y a escuchar..."* (74). Estos espacios son acompañados por un educador adulto.

A estas dos modalidades principales de organización de la jornada escolar, se agregan otros tres tipos de instancias educativas:

a. **Instancias de aprendizaje "incidental" y "sobre-incidental"**, en las que se aprovecha formativamente las situaciones emergentes en la vida cotidiana de la escuela (intereses e interrogantes expresados en forma circunstancial, conflictos entre estudiantes, sucesos de distinto tipo que sirven de punto de partida para la reflexión personal o colectiva), en especial en cuestiones o situaciones reales particularmente significativas por sus implicancias éticas o vinculadas a la convivencia.
b. **Instancias de aprendizaje "contextual"**, que se derivan del contacto con el mundo natural o cultural. Con el fin de propiciar que el contexto físico y cultural sea una fuente de aprendizaje permanente, la escuela debe proveer, por un lado, un ambiente rico en posibilidades experienciales y, por otro lado, oportunidades de encuentro con *"emisores culturales especializados"* externos a la propia institución: la familia, en primer lugar, pero también artistas, deportistas, científicos, otras instituciones especializadas en temas específicos. *"En una sociedad que transita hacia el gestionismo el desarrollo proyectivo de los jóvenes no puede resolverse en su totalidad desde las instituciones educativas generales. No son pocos -y deberían ser más- los casos de niños y jóvenes que manifiestan un particular interés, generalmente -aunque no siempre- unido a una particular aptitud natural, por tal o cual actividad, artística, científica, deportiva... El ambiente de la institución educativa proyectiva debe ser abierto a la relación con otros factores educativos... y emisores culturales"* (78-79).
c. **Instancias de "reflexión"**: Se trata de espacios formativos que cuentan con el acompañamiento de un educador adulto, en los que los estudiantes reflexionan y evalúan la marcha de sus proyectos, el modo en que están gestionando sus tiempos, las opciones

es necesario que existan otros de manejo personal sin los cuales no hay actividad proyectiva posible. El manejo de tiempos y espacios personales es una experiencia invalorable cuando se pretende el desarrollo de espíritus creativos y es uno de los primeros elementos a tener en cuenta en un ambiente pedagógico proyectivo" (76).

Con respecto a la evaluación, que es el tema que nos ocupa en este libro, se señala que *"en una institución educativa proyectiva en tanto emisor, se evalúan a más de la adquisición de conocimientos y la efectividad de la actividad docente, el desarrollo del niño o el joven en tanto sujetos de proyecto y la propia estrategia institucional. La evaluación de una institución proyectiva no se limita pues a cotejar resultados en relación a objetivos predeterminados por un emisor externo como sucede con los agentes educativos. En relación al carácter de la evaluación, ella es cualitativa y personal respecto al crecimiento del alumno, en tanto atiende a puntos de partida personales y de llegada explícita o implícitamente propios del sujeto. Y es en gran parte cuantitativa e impersonal respecto a la adquisición de conocimientos, particularmente de los conocimientos instrumentales"* (81).

La propuesta de educación proyectiva no utiliza calificaciones ni repetición de cursos. *"Nuestra propuesta elimina calificaciones y juicios sintéticos como modo de evaluación. Evaluamos al niño permanentemente en reuniones periódicas de docentes y de éstos con los coordinadores. Los padres reciben esta evaluación por parte de la coordinación en reuniones individuales periódicas"* (147). Estas evaluaciones se apoyan principalmente en reportes elaborados por los distintos educadores a partir de la observación del desempeño y actitudes de los niños (148).

Desde nuestra perspectiva, el aspecto de la evaluación no está suficientemente desarrollado en la propuesta de educación proyectiva. De todas formas, nuestro interés es mostrar la necesidad de un modelo escolar alternativo, como marco para plasmar una propuesta de evaluación superadora de las limitaciones propias de los sistemas de calificación vigentes.

Creemos que tres podrían ser las notas principales de una propuesta de evaluación en un modelo educativo alternativo como el presentado en las páginas previas.

En primer lugar, la evaluación formativa, tal como fue presentada en el capítulo anterior, se articula perfectamente con el desarrollo de la capacidad de proyectar. Concebir y realizar proyectos van de la mano con la capacidad para evaluar en el sentido formativo del término, incluyendo la autoevaluación y la coevaluación. La evaluación como forma de conocimiento, según fue explicada en el capítulo 1, es una parte esencial de la capacidad proyectiva de los seres humanos.

En segundo lugar, la capacidad proyectiva conceptualizada en las páginas anteriores, admite ser expresada bajo la forma de rúbricas que contemplen sus dimensiones principales en términos de actitudes, valores, saberes y habilidades (y de un modo que se adecuen a las características específicas de la actividad proyectiva en las distintas edades, así como a su progresión). La definición de las características centrales de los diversos productos y desempeños vinculados a la capacidad de proyectar, que cada estudiante debería poder realizar y demostrar a lo largo de su trayectoria escolar, permitiría romper con el modelo actual de calificaciones. La certificación no estaría sujeta a plazos predeterminados ni a la aprobación de asignaturas específicas. El estudiante debería demostrar su capacidad para concebir y llevar adelante un proyecto, en un área de su elección, que podría involucrar conocimientos y habilidades de diferentes disciplinas. Podría presentar su proyecto en el momento en que esté listo para hacerlo, con plazos flexibles, bajo la forma de "exhibiciones" o "desempeños de comprensión", los que serían valorados con criterios explícitos y conocidos con antelación. Tendría, además, oportunidades para rehacer y volver a presentar sus proyectos si no resultan adecuados.

Un enfoque de este tipo permitiría una enorme flexibilidad y diversidad en las formas, contenidos y tiempos de acreditación de los saberes logrados, incluyendo el reconocimiento formal de saberes adquiridos fuera del sistema educativo formal. Obviamente, en un modelo escolar como este, la repetición de grado no se plantea, sencillamente porque no tiene razones para existir (con la excepción, tal como plantea Ferrer (2002: 136), de situaciones excepcionales vinculadas con la maduración del niño o en casos de no asistencia o abandono de los cursos).

En tercer lugar, un tipo de evaluación complementaria en un sistema educativo alternativo como el propuesto, podría ser la realización de pruebas nacionales que se concentren en algunos pocos saberes fundamentales y comunes para todos, vinculados principalmente a la lectura, la escritura, el razonamiento lógico-matemático aplicado, así como algunos saberes científicos y culturales de especial relevancia. En la educación primaria y media básica estas pruebas podrían ser de carácter formativo, es decir, sin consecuencias para los estudiantes, pero orientativas para los centros de enseñanza. En la enseñanza media superior podrían ser parte del sistema de certificación.

5. ¿Qué hacer con las calificaciones? Mientras el futuro nos alcanza...[12]

En el apartado anterior expusimos la necesidad de concebir un sistema de evaluación para la certificación radicalmente diferente de los sistemas de calificaciones que tenemos actualmente. Ello requeriría de un modelo educativo alternativo, concebido y estructurado sobre otras bases, tanto en relación a sus propósitos formativos como en cuanto a sus formas organizativas. Presentamos sintéticamente y en forma ilustrativa, la propuesta de la educación proyectiva, que permitiría llevar adelante ese desafío.

En este apartado retornamos a la realidad que tenemos hoy: sistemas educativos organizados en grados y materias, currículos centrados en contenidos que deben ser enseñados secuencialmente, y que nos exigen a los docentes establecer una calificación para cada estudiante cada cierto tiempo. Nuestra intención en las páginas que siguen es proponer un conjunto de criterios y pistas concretas para realizar esa tarea de la manera más sensata posible, de modo que nuestras calificaciones tengan sentido para los estudiantes y no se conviertan en un obstáculo para el aprendizaje y para sus trayectorias educativas.

5.1. Dar notas con menos frecuencia

La primera sugerencia fundamental es reducir la frecuencia con que damos notas a los estudiantes, esto es, evitar calificar en forma diaria o semanal y evitar asignar calificaciones a todos los trabajos o actividades que desarrollan los estudiantes. En lugar de ello, es necesario desarrollar un esfuerzo sistemático por realizar devoluciones significativas, en el sentido explicado en el capítulo anterior. El principal esfuerzo que necesitamos hacer es romper con el hábito de asignar calificaciones o puntajes continuamente a los estudiantes, lo que, según vimos, es práctica habitual en buena parte de los docentes de la región.

La primera objeción a esta sugerencia seguramente será que familias y estudiantes "piden" las notas. Es una cuestión cultural. El sistema los ha acostumbrado a ello, necesitamos ahora des-acostumbrarlos y crear nuevas costumbres y una nueva cultura. Como es obvio, es difícil hacerlo en solitario,

12. Este subtítulo está inspirado en el título del libro de Margarita Zorrilla, reconocida maestra y experta mexicana, **Hacer visibles buenas prácticas. Mientras el debate pedagógico nos alcanza**. México: Consejo Mexicano de Investigación Educativa, 2005.

debería ser un esfuerzo colectivo, al menos a nivel de un centro de enseñanza o del equipo de docentes de una misma asignatura. Por otra parte, lo que las familias realmente demandan es información sobre cómo va el aprendizaje de sus hijos. Somos nosotros quienes respondemos a esa demanda con una nota como respuesta. Es cierto también que puede haber requerimientos administrativos del propio establecimiento o sistema educativo, en cuanto a que se debe registrar una cierta cantidad de calificaciones en cada período. Al respecto solo cabe decir que los docentes estamos acostumbrados a que hay muchas maneras de "cumplir" con los requisitos administrativos y, al mismo tiempo, mantener cierto margen de autonomía. Por cierto, lo deseable sería que las autoridades educativas trabajen en el desarrollo de sistemas de evaluación diferentes, alineados con los requerimientos de un modelo de educación alternativo.

Esta primera sugerencia implica recurrir a las múltiples formas de realizar devoluciones centradas en la tarea, analizadas en el capítulo anterior, bajo la forma de comentarios del docente, del uso de rúbricas o de instancias de co-evaluación o auto-evaluación, entre otras. Lo importante es enfocar al estudiante en la revisión y reflexión acerca del producto o desempeño en el que trabajó porque, según mostramos en el capítulo anterior, una vez que la nota aparece en escena desaparece el aprendizaje como foco de interés para la mayoría de los estudiantes.

5.2. Separar las instancias de certificación de las de evaluación formativa

De la mano con la primera recomendación va una segunda: es necesario separar las instancias de calificación de las de evaluación formativa (o, al menos, realizar algunas experiencias de este tipo a lo largo del curso). Además de evitar utilizar la calificación como medio de devolución habitual, necesitamos establecer instancias formales de evaluación que los estudiantes sepan que no serán calificadas, que su propósito será de carácter formativo, que tendrán la oportunidad de "equivocarse" y, sobre todo, que recibirán una devolución que les ayudará a rehacer y mejorar su trabajo o desempeño.

Para que esta estrategia funcione es importante que los estudiantes perciban el alineamiento entre las instancias de evaluación formativa y la futura "prueba" o instancia de evaluación para la certificación. En otras palabras, que perciban coherencia en el curso y que "aprender" de las evaluaciones formativas tendrá impacto sobre su desempeño en la evaluación que será calificada.

Con el fin de ilustrar esta propuesta relataremos el esquema de evaluación que los autores desarrollamos en el marco de un Diploma de Evaluación de Aprendizajes dirigido a docentes[13]. El Diploma tiene carácter semipresencial, con un encuentro mensual y actividades virtuales que los participantes deben realizar cada semana. El currículo está estructurado en módulos bimestrales. Para cada módulo se establece, desde el inicio del mismo y como parte del Programa de estudio, la forma que tendrá la evaluación para la certificación. Esta consiste en un trabajo final que denominamos "Trabajo de Acreditación del Módulo" (y que terminó siendo referido como "TAM"). Es decir que, al inicio de cada módulo, los participantes saben en qué consistirá el TAM y con qué criterios será valorado, lo cual está expresado en una rúbrica que es parte del Programa. A lo largo del módulo los estudiantes deben realizar una serie de actividades virtuales y subirlas a la plataforma del curso. Estas actividades son evaluadas por docentes tutores, pero no calificadas. Es importante señalar que no solo no son calificadas en forma abierta, sino que se garantiza a los estudiantes que tampoco se llevará una "contabilidad oculta" o registro de notas que ellos no conozcan. Cada estudiante recibe una devolución a su trabajo por parte del docente tutor y, en muchos casos, a través de Foros, por parte de sus compañeros. Lo importante de esta forma de proceder es que:

a. los estudiantes, al conocer desde el comienzo cómo será el trabajo de acreditación, pueden percibir el vínculo entre las actividades formativas y lo que se les pedirá en la evaluación final;
b. esto da importancia a las actividades formativas;
c. como las actividades virtuales no son consideradas para la calificación, los participantes tienen "libertad para equivocarse", para presentar productos en borrador, inacabados o parcialmente desarrollados; esto es muy importante en un contexto en que la posibilidad de ser calificado a veces hace que los estudiantes no se "arriesguen" a entregar producciones incompletas o que no los satisfacen plenamente. Nuestra consigna era *'más vale presentar algo, aunque no esté completo, porque eso habilita a que haya devoluciones y sugerencias por parte del docente y de los colegas'*.

De todas formas es importante notar que la viabilidad de este tipo de estrategia depende, en parte, de las edades de los estudiantes y de la frecuencia de las clases. El planteamiento anterior es particularmente válido para el

13. El Diploma en Evaluación de Aprendizajes lo iniciamos en el Instituto de Evaluación Educativa de la Universidad Católica del Uruguay en el 2009 y seguía funcionando en el momento de escribir este libro (2016).

nivel superior, en que la frecuencia de clases es baja y la autonomía de trabajo de los estudiantes es alta. En los niveles primario y medio puede resultar conveniente llevar un registro más frecuente de los avances y logros de los estudiantes, con el fin de tomarlos en consideración a la hora de calificar. Es razonable suponer que, en estos niveles, la calificación esté más vinculada a procesos de trabajo cotidianos que a un único producto final. En cualquier caso, sigue siendo importante evitar usar calificaciones como forma de "devolución" y enfocarse más en el registro y análisis de las características de los desempeños y tareas específicas realizadas por los estudiantes. Veremos un ejemplo un poco más adelante.

5.3. Informar y clarificar los principales aprendizajes o logros que serán tenidos en cuenta para la calificación

Esta recomendación es una reiteración de lo expresado en el capítulo anterior acerca de la importancia de *"clarificar y compartir las intenciones educativas y criterios de logro"*. Es el complemento de la misma, pero desde la perspectiva de la evaluación para la certificación. Tal como fue indicado, no se trata simplemente de escribir objetivos y metas de aprendizaje sino, sobre todo, de desarrollar el "olfato" de los estudiantes para reconocer las características de los desempeños de calidad en la materia de que se trate.

Los docentes solemos pensar mucho en los contenidos que vamos a enseñar y poco en los aprendizajes que pretendemos que los estudiantes construyan. Normalmente, la mayoría de los docentes hacemos a los estudiantes una breve presentación de los contenidos que vamos a trabajar en un curso o unidad. Raramente explicamos qué aprendizajes esperamos que logren al final del mismo. Más raro aún es que expliquemos desde el comienzo del curso en qué consistirán específicamente las actividades de evaluación para la certificación. A lo sumo indicamos qué cosas llevarán nota y qué peso tendrá cada una en la nota final. Pero el tipo de saberes y habilidades que serán parte de la prueba o examen final, así como el significado de los niveles aceptables, suelen ser una sorpresa hasta el momento final (para algunos estudiantes nunca dejan de ser un misterio).

Tal como fue explicado en el capítulo anterior con relación a la definición de las dimensiones de una rúbrica, es importante identificar y expresar con claridad cuáles son los aspectos centrales y relevantes de las intenciones educativas del curso o unidad, de un modo en que puedan diferenciarse claramente entre sí. Dicho con otras palabras, cuáles son los principales aprendizajes (saberes y habilidades), relacionados con el contenido de la materia, que los estudiantes deberán demostrar al final del curso o unidad y que serán

tenidos en cuenta para calificar su desempeño. Es importante especificarlos tanto en términos de "dimensiones", como en términos de niveles progresivos de logro, al estilo de lo planteado con relación a las rúbricas.

5.4. Valorar y reportar por separado cada dimensión relevante de nuestras intenciones educativas

Una práctica de enorme importancia, que se deriva con claridad de los análisis realizados a lo largo de este capítulo, consiste en calificar por separado cada dimensión relevante, evitando, en lo posible, el uso de promedios matemáticos. En caso de que reglamentariamente sea necesario llegar a una calificación única -lo cual, según vimos antes, suele ser la norma y no una excepción- es conveniente establecer la práctica de reportar a estudiantes y familias las calificaciones desagregadas que están en la base de la calificación "resumen", de modo que puedan comprender el modo en que esta fue construida y su significado en términos de qué dimensiones fueron las más y las menos logradas.

Para entender mejor este aspecto es útil pensar lo que ocurre con ciertos desempeños artísticos o deportivos que son objeto de evaluación por jueces, tales como un conjunto de carnaval, el nado sincronizado o la gimnasia olímpica. En todos estos casos el desempeño es evaluado por un jurado integrado por varios miembros, para evitar la probable y casi inevitable arbitrariedad de un único juez. Cada uno de los miembros del jurado observa y valora diversos aspectos de la performance de los competidores. Por ejemplo, en el caso de la gimnasia olímpica, el ritmo, la sincronización con la música, cierto tipo de saltos o movimientos, el modo en que el atleta "cae" al final de una pirueta, entre otros. De un modo similar funcionan los jurados de nado sincronizado, de concursos artísticos y de muchas otras disciplinas. Los puntajes otorgados a los distintos aspectos relevantes de la performance (que equivalen a las "dimensiones" de una rúbrica), se promedian para establecer un puntaje único, que determina el orden de los competidores y, como consecuencia de ello, quiénes accederán a las medallas o a los premios del concurso.

En estos ejemplos, la única utilidad del puntaje promedio final es ordenar a los participantes en el marco de la competencia. Pero si alguno de los participantes desea aprender algo de esta experiencia y tener alguna pista para mejorar, el promedio global no le servirá de mucho. Necesitará acceder a los puntajes otorgados a cada aspecto relevante, para saber en cuáles fue bien evaluado y en cuáles no.

Algo similar ocurre en la educación. La calificación global final sirve únicamente para definir la aprobación o no del curso (asumiendo que ya no estamos interesados en propiciar la competencia entre los estudiantes ni en distinguir entre mejores y peores). Pero solo la calificación desagregada por dimensiones puede ofrecer alguna información útil para futuros desempeños o para comprender el significado de la calificación global.

Una pista concreta sobre cómo llevar esto a la práctica la propone Wiliam (2011: 124), para una clase de ciencias. Los docentes establecieron las siguientes diez dimensiones o áreas de interés para un bimestre:

a. Uso apropiado del equipamiento de laboratorio
b. Conversión y etiquetado de unidades de medida
c. Cálculo de densidad
d. Aplicaciones del concepto de densidad (flotación, hundimiento, estratificación, expansión térmica)
e. La densidad como propiedad característica
f. Fases de la materia (a nivel molecular)
g. Leyes de los gases
h. Comunicación (graficar)
i. Comunicación (elaborar informes de laboratorio)
j. Habilidades para la investigación

Para cada una de estas dimensiones, se identificaron fuentes de evidencia: observación en el aula, tareas domiciliarias, informes escritos por los estudiantes, entre otros. Sobre la base de dichas evidencias los docentes registraban en una planilla su valoración para cada estudiante en cada dimensión, utilizando tres categorías muy sencillas: 0 (no hay evidencia de dominio en esta dimensión); 1 (hay alguna evidencia de dominio); 2 (hay fuerte evidencia de dominio). En la planilla mencionada cada línea corresponde a un estudiante, cada columna a una dimensión y en cada celda se consigna la valoración correspondiente (0, 1 o 2). A cada celda se le agrega además un color específico según su valor (por ejemplo, rojo para 0, amarillo para 1, verde para 2). El resultado es una suerte de "mapa" de la clase, que permite en forma rápida observar qué ocurre con cada dimensión y qué ocurre con cada estudiante. En particular, es posible observar a simple vista qué le está faltando a cada estudiante.

"Al final del período, los estudiantes rinden una prueba, que es utilizada para confirmar la evidencia recogida hasta ese momento. Cuando el desempeño en la prueba indica un nivel de dominio diferente del que surge de la planilla de registro del trabajo en clase a lo largo del período, esto es considerado como algo que requiere mayor indagación por parte del docente (en lugar de simplemente hacer un promedio entre la prueba y los registros). Típicamente el docente entre-

vistará al estudiante para profundizar el análisis del grado de comprensión del estudiante sobre los aspectos en cuestión" (Wiliam, 2011: 125).

La aplicación práctica de esta forma de calificar en grupos de 8º grado mostró que obliga a estudiantes y docentes a pensar en plazos más largos, dado que no hay una nota para cada tarea o actividad, sino un registro que será confrontado con los resultados de una prueba al final del período, como resultado de lo cual se definirá la calificación. Los estudiantes se mostraron más involucrados en monitorear su propio aprendizaje y progreso, pedían clarificación con mayor frecuencia y veían al docente más como un apoyo que como un juez.

También es importante mantener por separado, en caso de que vayan a formar parte de la calificación final, las valoraciones de aspectos tales como el comportamiento, actitudes, prolijidad, "esfuerzo", etc., evitando hacer promedios de cosas disímiles entre sí. Según vimos antes en este capítulo, esta es una práctica bastante extendida. Es necesario entonces definir con claridad cuál es el significado específico de cada uno de dichos aspectos, cómo se lo valorará y de qué forma integrará la calificación. Para algunos de estos aspectos, los más importantes, sería deseable desarrollar una rúbrica que clarifique a los estudiantes lo que se espera de ellos y que sirva de encuadre a la subjetividad del docente, para evitar que devenga en arbitrariedad.

5.5. Explicitar las fuentes de evidencia y/o actividades que se utilizarán en la evaluación de certificación desde el comienzo del bimestre, unidad didáctica o módulo curricular

Junto con conocer y comprender los aprendizajes que deben lograr a lo largo de un curso o unidad, los estudiantes deberían conocer y comprender también desde el inicio qué tipos de tareas o productos deberán realizar para la certificación, en qué momentos tendrán lugar, con qué criterios serán valorados y qué otras evidencias utilizará el docente para calificarlos (pruebas, proyectos, informes, exposiciones u otros productos o formas de desempeño).

Definir y explicitar estos elementos desde el comienzo del curso puede resultar enriquecedor para el docente, porque le permite asegurarse que exista coherencia entre el enfoque de enseñanza del curso y la evaluación. Habitualmente los docentes planificamos el curso primero, lo damos luego, y recién al final pensamos en cómo realizar la evaluación con fines de calificación. En los últimos años se ha desarrollado un enfoque de planificación denominado "diseño retrospectivo" (Wiggins y Mc Tighe, 2005), que consiste en comenzar la planificación por el final, es decir, por la evaluación.

Una vez diseñada en términos generales la evaluación final de una unidad, el docente planifica la enseñanza de modo tal de asegurarse que los estudiantes tendrán oportunidades para aprender lo que se necesita para desempeñarse adecuadamente en la evaluación.

Es importante también explicitar el enfoque desde el cual se realizarán las valoraciones del desempeño, sobre todo en los casos en que se construye la calificación a partir de la constatación del avance o progreso de cada estudiante con respeto a su situación en un momento previo. Según vimos en algunos testimonios aportados antes en este capítulo, este es un enfoque que adoptan muchos docentes. Lo analizaremos con más detalle en el apartado siguiente.

5.6. Dar significado sustantivo a los distintos tramos de la escala de calificaciones

Este aspecto está implícito en los apartados anteriores, pero conviene explicitarlo con énfasis. Independientemente de cuál sea la escala de calificaciones vigente en el sistema educativo en el que uno trabaja, es importante hacer el esfuerzo de expresar el significado de los principales tramos de la escala (en no más de cuatro o cinco tramos), en términos de desempeños. Lo recomendable es organizar la escala en torno a cuatro grandes categorías, del tipo de las indicadas en la Figura 4.5 del capítulo anterior, para luego construir una rúbrica que explicite el significado sustantivo de cada tramo. Al interior de estos tramos uno puede luego establecer matices, según qué tan amplia sea la escala. Lo importante es evitar que la calificación sea el resultado de operaciones matemáticas sin significado sustantivo.

Imaginemos por ejemplo que la escala de calificaciones va de 1 a 12 y que 6 es la nota de suficiencia. Uno podría establecer los siguientes cuatro tramos: 9-10 significa destacado (más allá de lo esperado); 7-8 significa que el estudiante ha logrado plenamente lo esperado; 6 que está en el límite de lo aceptable; 3-5 que está en proceso hacia un nivel aceptable. Uno debería luego describir las características de los desempeños y/o producciones necesarias para ser ubicado en cada uno de estos cuatro tramos. La nota todavía no entra en juego. Una vez que se ha ubicado el desempeño del estudiantes dentro de un tramo, uno puede pasar a "refinar" la valoración asignando una de las notas correspondientes al tramo en cuestión. En la Figura 5.5 incluimos un ejemplo que puede ayudar a comprender la idea planteada.

Figura 5.5: Ejemplo de rúbrica para calificar ensayos en un colegio secundario de Colorado

9-8	Las respuestas en el rango superior satisfacen los siguientes criterios: *Resumen*. El resumen debe identificar la idea central [del material leído] *Foco del acuerdo*. El acuerdo y/o desacuerdo con el texto puede ser total o parcial, pero el estudiante debe dejar en claro con qué está de acuerdo y/o en desacuerdo. Específicamente, los trabajos en el rango 9-8 deben enfocarse en la tesis del autor, no en el abuso de sustancias en general. *Sustento de los acuerdos y/o desacuerdos*. El sustento debe proporcionar un análisis de los argumentos y/o ejemplos relevantes y concretos. *Estilo y coherencia*. Estos ensayos deben demostrar estilo claro, buena organización general y pensamiento consecuente. Deben tener pocos errores gramaticales repetidos. [Las cuatro frases en itálica representan las dimensiones que están siendo calificadas]
7	Esta calificación es utilizada para trabajos que cumplen con los requerimientos básicos para las calificaciones 9-8 pero tienen menor desarrollo, sustento o análisis.
6-5	Los trabajos en el rango medio omiten o tienen insuficiencias en alguno de los siguientes cuatro criterios: *Resumen*. No hay resumen o está incompleto, limitándose a enumerar las tesis del autor. *Foco del acuerdo/desacuerdo*. No está claro con qué está acordando y/o discrepando el estudiante o no tiene relación con lo propuesto por el autor. *Sustento de los acuerdos y/o desacuerdos*. El estudiante solamente se opone al autor sin argumentar; los ejemplos son muy generales o no se diferencian de los ejemplos incluidos en el artículo. El análisis es engañoso, irrelevante o estrecho. *Estilo y coherencia*. Estos ensayos tienen una organización muy débil o contienen errores gramaticales notorios.
4	Esta calificación es utilizada para trabajos un poco más débiles que los 6-5. Además, un estudiante que escribe su propio ensayo paralelo con un estilo competente recibirá un 4.
3-2	Estos trabajos son deficientes en *dos* o más de las dimensiones. Típicamente parafrasean el artículo o tienen problemas serios de organización y coherencia. Los trabajos con errores gramaticales serios y reiterados deben ser ubicados en esta categoría. [Todo este párrafo, al igual que los anteriores, es un descriptor para este tramo de la escala de calificaciones]

Fuente: Traducido y adaptado de Wiggins (1998: 155).

Siguiendo y, obviamente, adaptando este ejemplo, a lo largo del curso deberíamos explicitar e ilustrar para los estudiantes qué es necesario hacer o demostrar para acceder a los distintos niveles de calificación (no a cada punto de la escala, sino cada uno de los grandes tramos). Es decir, qué características deberá tener un producto o desempeño para ser considerado como en proceso, suficiente, logrado o destacado. Estas definiciones deberían ser ejemplificadas con trabajos o desempeños típicos, realizados por otros estudiantes en otros momentos.

En este punto es conveniente reiterar la reflexión acerca de la cuestión de la subjetividad en las calificaciones. La evaluación siempre tiene un componente de subjetividad, aun cuando se lo intente neutralizar a través de la asignación de puntajes y la matematización de las notas (en este último caso la subjetividad opera en el momento de construir las actividades y decidir qué puntaje tendrá cada una). En la atribución de valor a los productos o desempeños de los estudiantes siempre existe un componente de apreciación subjetiva por parte del docente. La diferencia entre subjetividad y arbitrariedad radica en el carácter explícito o implícito de los criterios de valoración. Cuando un docente califica a partir de "todo lo que sabe" de un alumno, sin explicitar qué dimensiones está considerando y cómo las valora, está siendo arbitrario. Cuando el docente realiza valoraciones a partir de criterios explícitos, conocidos y comprendidos por los estudiantes y, por tanto, abiertos a la discusión, la subjetividad sigue estando presente pero de un modo no arbitrario.

5.7. Ofrecer segundas oportunidades para los trabajos y pruebas vinculados a la certificación

Una última sugerencia para finalizar. Uno de los problemas de las instancias formales de evaluación para la certificación (pruebas, exámenes, parciales, entrega de trabajos) es que el estudiante tiene una única oportunidad para presentar su trabajo o desempeño. Esto no suele ser así en la mayor parte de las instancias de trabajo en la vida real, fuera del sistema educativo formal. Si bien hay actividades específicas en las que no existe una "segunda oportunidad" como, por ejemplo, despegar un avión o realizar una cirugía, en la mayor parte de las actividades de producción material e intelectual, las personas contamos con instancias de revisión de nuestra producción inicial, para posteriormente mejorarla.

Esta es una práctica que podemos incorporar a nuestros cursos, ofreciendo siempre al menos dos instancias de evaluación formal. Una o varias de carácter formativo en las que, en lugar de una calificación, los estudiantes recibirían devoluciones del docente y/o de sus compañeros, junto con orientaciones sobre qué es necesario mejorar para avanzar hacia un mayor nivel de logro. Luego vendría una instancia final, que será la que reciba una calificación.

Esto es relativamente sencillo de hacer cuando se evalúa a través de un ensayo, un informe, un proyecto de investigación, una producción escrita o alguna otra forma de producción parecida. Es un poco más complejo para aquellos docentes y materias que utilizan pruebas estructuradas en torno

a "ejercicios" más o menos breves. En estos casos, obviamente, la segunda instancia no podría estar constituida por los mismos ejercicios. Por tanto, implicaría para el docente elaborar dos pruebas enfocadas en los mismos conceptos, principios o habilidades. Por cierto, esto implica más trabajo, pero redundaría en mayor comprensión por parte de los estudiantes acerca de qué está siendo evaluado y, probablemente, en mayor cuidado en la confección de la prueba por parte del docente (muchas veces las construimos como conjunto de ejercicios relacionados con los principales temas de la unidad, pero sin una clara explicitación, incluso para nosotros mismos, de qué desempeños queremos evaluar). Aquellos docentes que suelen realizar una prueba de este tipo cada mes, podrían pensar en realizarla un mes con carácter formativo y el mes siguiente con fines de calificación.

La principal ventaja de introducir una práctica de este tipo es habituar a los estudiantes a analizar su propio desempeño, con el incentivo de que tendrán la oportunidad de mejorarlo, en lugar de quedarse con la nota como una valoración de su trabajo que tiene algo de azaroso o arbitrario, que es poco comprensible y sobre la cual no tienen nada más qué hacer.

6. Tres enfoques principales en la formulación de juicios de valor

Para cerrar este capítulo consideramos de interés explicar los tres enfoques principales para la definición de juicios de valor en los procesos de evaluación. Cada uno de estos enfoques tiene una lógica propia y los docentes solemos utilizarlos en forma combinada, en general sin tener del todo claras sus implicancias.

Siempre que calificamos estamos realizando un juicio de valor sobre los procesos y/o los logros de un estudiante a lo largo de un período de tiempo. Este juicio de valor se puede expresar de diversos modos, pero su expresión formal es la escala de calificaciones vigente en el sistema educativo en el que estamos. Para establecer el juicio de valor nos basamos en distintas formas de evidencia empírica sobre el desempeño del estudiante: lo que hizo en una prueba, lo que observamos acerca de sus actitudes y procesos cognitivos en el aula, las tareas que presentó, entre otras.

Ahora bien, es fundamental darse cuenta de que, a la hora de usar esas evidencias para construir un juicio de valor, siempre comparamos lo que emerge de las mismas con un punto de referencia. Esta es la esencia de la valoración: contrastar lo que "hace" el estudiante con algún referente o criterio de valor. Los principales puntos de referencia suelen ser tres: el desempeño anterior del propio estudiante, el desempeño de los otros estudiantes o una

conceptualización del desempeño esperado. Cada uno de estos puntos de referencia está en la base de un enfoque de evaluación.

6.1. El enfoque "normativo"

Un primer modo de establecer juicios de valor consiste en comparar el desempeño del estudiante con el de los demás de su grupo. Esto permite establecer quiénes lo han hecho mejor y quiénes peor. Lamentablemente, como ya dijimos, esto muchas veces se transforma en una valoración global de los individuos: los "mejores" y "peores" estudiantes.

A este enfoque se lo denomina **normativo**, por su relación con la "curva normal". El procedimiento formal consiste en ordenar a los estudiantes en función de sus puntajes, lo que da lugar a una distribución tipo "curva normal". En dicha curva una parte de los estudiantes -la mayoría- queda ubicado en el entorno del promedio, algunos se destacan hacia arriba (los mejores) y otros hacia abajo (los peores). En sus orígenes este enfoque estuvo vinculado a la elaboración de pruebas cuyo propósito era la selección de individuos para el ingreso a las universidades. En estas pruebas no importaba tanto qué es lo que los estudiantes específicamente "sabían" o dominaban, sino la distinción entre "mejores" y "peores" y su ordenamiento con la mayor "precisión" posible.

Esta práctica se trasladó a las calificaciones en el aula. Todos los docentes tendemos a establecer comparaciones al interior de un grupo, observando a veces matices muy específicos, con el fin de ser "justos" y "premiar" con una nota un poco más alta el trabajo que está un poco mejor que otro o al estudiante que se esforzó más que otros. Este "afán de justicia" está en la base de la amplitud de las escalas de calificaciones que, además, a muchos docentes nunca les resulta suficiente. En Uruguay, por ejemplo, en que hay 12 puntos en la escala, muchos docentes tienden a usar un '7 para arriba' o 'un 6 para abajo' como notas intermedias. Nos cuesta mucho trabajar con 4 o 5 categorías básicas, por este énfasis en la comparación <u>entre</u> estudiantes. Algo similar ocurre, según vimos, en países en los que las notas incluyen números decimales.

Tal como fue dicho, este enfoque tiene dos consecuencias perversas: promueve la competencia entre los estudiantes por la nota (entre aquellos a los que les va bien) y se transforma en una valoración de la persona que inmoviliza (entre aquellos a los que no les va bien).

6.2. El enfoque "criterial"

Un segundo modo de establecer juicios de valor consiste en comparar el desempeño del estudiante contra un "criterio" (de allí la denominación) o definición de lo que se espera que logre. Este criterio o definición puede ser explícito o implícito. Normalmente los docentes tenemos interiorizada una imagen de lo que esperamos que los estudiantes realicen o logren: qué es un buen escrito, qué actitudes en clase consideramos positivas, qué tipo de comportamientos son deseables, etc. Detrás de cada aspecto que tenemos en cuenta a la hora de calificar, hay una cierta imagen de lo que esperamos. El problema fundamental, según hemos visto antes, es que nos cuesta mucho (porque no es tarea fácil) transparentar, explicitar, hacer comprensibles para los estudiantes estas imágenes o expectativas. A esto es a lo que nos hemos referido al hacer mención a la necesidad de clarificar las intenciones educativas.

La falta de explicitación de estos criterios que actúan como referentes en la evaluación es, como vimos antes, una de las principales debilidades de los diseños curriculares y los reglamentos de evaluación. Es el eslabón perdido entre ambos.

La evaluación desde un enfoque criterial requiere de la formulación clara de las intenciones educativas en una determinada materia y unidad, identificando claramente las dimensiones y niveles, tal como explicamos para la construcción de una buena rúbrica. Esto permite establecer un punto de referencia claro contra el cual contrastar la evidencia acerca de los procesos y logros de los estudiantes.

En este caso, la nota no surge de una comparación entre estudiantes, sino de una comparación entre el desempeño real del estudiante y el desempeño esperado y explicitado.

6.3. El enfoque de "progreso"

Un tercer modo de establecer juicios de valor consiste en comparar el desempeño del estudiante con su propio desempeño en un momento anterior en el tiempo. Este enfoque suele denominarse de "progreso" o de "crecimiento". Según vimos antes, a los docentes que utilizan este enfoque no les preocupa tanto qué logró el estudiante, sino en qué medida está avanzando con respecto a su punto de partida. Tal como advertimos antes, es importante establecer una distinción entre *progreso* y *esfuerzo*. Un estudiante puede realizar grandes esfuerzos sin progresar.

El enfoque de progreso requiere que uno registre el desempeño del estudiante en varios momentos del año, en un mismo tipo de tarea, conoci-

miento o habilidad, con el fin de constatar avances o mejoras. Esta aclaración es importante, porque a veces se producen ciertas confusiones. Por ejemplo, un docente podría afirmar que un estudiante "progresó" porque en el primer bimestre le fue mal en una prueba sobre fracciones y en el tercero le fue muy bien en una prueba sobre geometría. Esto no es progreso, porque el estudiante sigue sin comprender fracciones. Progreso implica volver a evaluar o registrar desempeños en un mismo tipo de saberes, con el fin de constatar la adquisición de habilidades, una comprensión más profunda del tema u otro tipo de cambios.

Hechas estas salvedades, es legítimo tomar en cuenta el progreso a la hora de calificar a los estudiantes. En este caso no importa comparar a los estudiantes entre sí ni con un desempeño ideal (si bien una cierta definición del mismo siempre estará subyacente). Lo que interesa comparar es la situación actual del estudiante con respecto a su situación anterior en el tiempo.

6.4. Los tres enfoques en el aula

Por lo general, al calificar a los estudiantes, la mayoría de los docentes empleamos alguna forma de combinación de estos tres enfoques. Por un lado, tenemos en cuenta qué alumnos han mostrado un nivel aceptable o satisfactorio con relación a nuestras intenciones educativas para el curso (a una imagen interna de la misma o a metas explícitas), y cuáles no –enfoque criterial–. Simultáneamente, es probable que muchos tengamos en cuenta el punto de partida de cada alumno -enfoque de progreso-. Por ejemplo, si un estudiante aún no alcanzó un nivel satisfactorio, pero había comenzado muy mal, se esforzó mucho y logró mejorar con respecto al momento inicial, es posible que le asignemos una calificación un poco más alta que la que correspondería por el enfoque criterial. En cambio, si otro estudiante tuvo un buen desempeño, pero nos consta que no se ha esforzado mucho y que en realidad tiene potencial para hacerlo mucho mejor, es posible que le asignemos una calificación inferior a la que le correspondería.

Paralelamente, los docentes solemos hacer comparaciones entre los estudiantes cuando definimos las calificaciones -enfoque normativo-. Por ejemplo, tomamos las cinco o seis mejores pruebas y las comparamos entre sí. A alguna le asignamos una calificación un poco más alta que a otra, porque si bien ambas pruebas son igualmente destacadas, una de ellas está un poco mejor. Ya mencionamos esta suerte de afán que tenemos los docentes por la "justicia calificativa".

Otro modo de proceder, dentro de un enfoque normativo, consiste en asignar la calificación más alta a los mejores trabajos, para luego calificar al

resto en forma relativa y proporcional a los mismos, "descendiendo" progresivamente. Este modo de proceder es especialmente riesgoso si es utilizado en forma exclusiva, sin emplear alguna definición o criterio de desempeño mínimamente aceptable. Cuando se califica a los estudiantes teniendo en cuenta únicamente la comparación de los desempeños dentro del grupo, si el nivel general del grupo es insuficiente, un alumno puede ser "el mejor del grupo", pero aún así no haber alcanzado un desempeño satisfactorio. Este es un serio problema de las calificaciones cuando no existe una buena descripción de los aprendizajes esperados: los docentes muchas veces nos adecuamos al "nivel" del grupo, reducimos nuestras expectativas y esperamos de los estudiantes menos de lo que serían capaces de lograr.

Por lo expuesto, el enfoque normativo aparece como el menos relevante para la evaluación en el aula. Es un enfoque cuya utilidad principal está vinculada a aquellas evaluaciones que tienen como finalidad el ordenamiento de individuos para procesos de selección como, por ejemplo, los concursos para el acceso a cargos de dirección.

El enfoque de progreso es especialmente importante en la educación básica. Tal como afirmamos antes, durante la educación inicial, primaria y secundaria básica, el foco debería estar puesto en evaluaciones y dispositivos de apoyo que ayuden a cada estudiante a avanzar en el desarrollo de sus capacidades, flexibilizando los tiempos para alcanzar los niveles de desempeño esperados y evitando la repetición de cursos. No obstante, la presencia de elementos propios del enfoque criterial es siempre importante. El progreso solo puede ser apreciado adecuadamente en la medida en que existen descripciones claras y explícitas de los aprendizajes a lograr.

Finalmente, es importante notar que el enfoque de progreso pierde valor en los niveles superiores del sistema educativo. En la educación terciaria, en particular, una de cuyas finalidades principales es otorgar títulos que habilitan al desempeño de diversas profesiones, la preeminencia del enfoque criterial es evidente. En estos niveles la certificación implica mucho más que aprobar un curso. Significa que el individuo tiene la preparación indispensable para el ejercicio profesional. Si no ha logrado esta preparación, no debería acceder a la certificación, aun cuando haya progresado mucho con respecto a su punto de partida.

7. Resumen y sugerencias para el trabajo en el centro educativo

A lo largo de este capítulo hemos analizado las normas y prácticas que rigen la valoración del trabajo de los estudiantes por parte de los docentes,

en los sistemas educativos tal como están actualmente organizados (con currículos organizados en materias y ciclos anuales que deben ser aprobados en tiempos predefinidos para poder seguir adelante). Vimos que los reglamentos de evaluación suelen asignar más importancia a las escalas de calificación y a los procedimientos para establecer una nota, que a los significados sustantivos de las mismas en términos de logros educativos. Esto es fruto de un desencuentro entre los reglamentos de evaluación y los diseños curriculares. Analizamos también la falta de sentido de la práctica de hacer que un estudiante vuelva a cursar un grado completo cuando se supone que ha logrado los aprendizajes esperados en alrededor de la mitad (a veces más) de las materias.

En este marco, aportamos evidencia recogida en nuestros trabajos de investigación en la región, que muestra que la mayoría de los docentes tendemos a calificar continuamente a los estudiantes. En el capítulo anterior habíamos analizado los efectos nocivos de esta práctica: las calificaciones se transforman en el objetivo del estudiante, el aprendizaje pasa a segundo plano, la nota se constituye en una herramienta de disciplinamiento más que de información sobre logros y, simultáneamente, se transforma en un "atributo" de la persona más que de su aprendizaje: fulano es un '5'. Inevitablemente, la nota desplaza a las devoluciones y a la evaluación formativa del escenario educativo.

Describimos además distintas formas típicas de llevar adelante la práctica de la calificación, que oscilan entre dos extremos: la valoración absolutamente subjetiva y arbitraria (*'yo tengo la imagen de que a fulano le tengo que poner tanto'*), por un lado, y la matematización de la nota (la nota es el resultado de un conjunto de operaciones matemáticas con puntajes que luego se promedian), por otro. Simultáneamente, en ambos extremos de este espectro, cada docente tiene su particular combinación de aspectos para incluir en la valoración del estudiante: pruebas, actitudes, esfuerzo, disciplina, tareas fuera de clase y un largo etcétera. El resultado es que el significado de las calificaciones es completamente opaco para estudiantes y familias. Nadie sabe a ciencia cierta qué significa una determinada nota en el sistema educativo, más allá de conocer su consecuencia: la aprobación o reprobación de un curso por parte del estudiante.

Vimos que estos problemas están estrechamente vinculados al modelo de escuela propio de la revolución industrial, la mayor parte de cuyos elementos constitutivos son obsoletos. Superar las limitaciones de los sistemas de calificaciones requiere redefinir el conjunto del modelo. El tratamiento en profundidad de este tema excede los alcances de este libro, pero dejamos

planteadas algunas ideas, tomadas básicamente de la propuesta de Educación Proyectiva de Ferrer.

"Mientras el futuro nos alcanza", es decir, mientras esperamos una transformación del ADN de los sistemas educativos propios de la sociedad industrial, los docentes debemos seguir adelante con nuestro trabajo. Con el fin de mejorar las formas en que calificamos a nuestros estudiantes en el marco del modelo predominante, propusimos un conjunto de pistas de trabajo que, en nuestra opinión y experiencia, permitirían utilizar mejor las calificaciones, evitar sus efectos perversos y ayudar a que los estudiantes comprendan cómo los evaluamos.

Sobre el final del capítulo introdujimos la distinción entre los enfoques normativo, criterial y de progreso en la construcción de las calificaciones.

Al igual que en los capítulos anteriores, cerramos este con algunas propuestas para trabajar en conjunto con algunos colegas en torno a los temas analizados.

Sugerencias para el trabajo en el centro educativo

La complejidad y dificultad de la asignación de calificaciones, cuando se hace con responsabilidad, determina que la cantidad de aspectos a revisar y trabajar sea muy amplia.

A continuación ofrecemos algunas sugerencias que podrían servir para iniciar un proceso junto con otros colegas, a lo largo de varios encuentros.

1. Un primer ejercicio, que puede resultar de utilidad para empezar a dialogar sobre el tema y conocerse un poco mejor, consiste en volver a leer los apartados 2 y 3 de este capítulo, preguntándose cada uno: ¿En qué docentes y prácticas de las presentadas me siento reflejado? ¿Qué aspectos tengo en cuenta habitualmente al calificar a mis estudiantes? ¿Qué procedimientos sigo para llegar a una calificación? ¿Hay algo que querría modificar? ¿Qué cosas y en qué sentido? Luego de un tiempo de lectura y reflexión individual, se puede tener un tiempo para compartir y escucharse.
2. Un segundo ejercicio consistiría en tener un intercambio en torno a pruebas empleadas para calificar a los estudiantes. Cada docente puede aportar una prueba y explicar cuál sería un desempeño aceptable en la misma y qué características tendría un desempeño plenamente satisfactorio. ¿Qué diferencias hay entre ambos?

3. Volver sobre el apartado 4.5. y construir una planilla como la propuesta por Wiliam, pero para una unidad de la materia que cada uno enseña. Revisarla con los colegas e intercambiar sugerencias. Utilizarla durante un período de uno o dos meses, tal como lo propone Wiliam. Registrar notas de la experiencia para compartirla con los colegas.

4. Construir una rúbrica al estilo de la Figura 5.5 para un producto específico que sirva como evaluación final de una unidad o semestre. Discutir la rúbrica con los colegas. Utilizarla para la corrección de los trabajos. Registrar notas de la experiencia para compartirla con los colegas. Rehacer la rúbrica a partir de la experiencia.

5. Elegir una unidad del curso y hacer la experiencia propuesta en el apartado 5.7, es decir, proponer dos instancias de evaluación, la primera de ellas sin calificación pero con devolución y con la oportunidad de rehacer y mejorar el primer intento. Explicitar a los estudiantes desde el comienzo cuáles son los aprendizajes a lograr, cómo se desarrollará la experiencia y qué se requerirá para alcanzar cada tramo de calificaciones.

Al igual que en el capítulo anterior, les sugerimos que cada docente lleve un registro escrito de sus reflexiones sobre la experiencia (lo que "funcionó" y lo que no, lo que aprendió y lo que debería modificar o mejorar), para luego compartirlo con los colegas.

Referencias bibliográficas

Ferrer, L. (2006). *Educación Proyectiva. Un aporte al debate educativo.* Montevideo: Tradinco.

Loureiro, G. (2009). *Evaluación en el aula, currículo y evaluaciones externas*. Lima: GTEE- PREAL.

Picaroni, B. (2009). *La evaluación en las aulas de primaria: usos formativos, calificaciones y comunicación con los padres*. Lima: GTEE-PREAL.

Ravela, P. (2006). *Fichas didácticas. Para comprender las evaluaciones educativas*. Santiago de Chile: PREAL.

Ravela, P. (2009). *Consignas, devoluciones y calificaciones: los problemas de la evaluación en las aulas de educación primaria en América Latina.* En: Páginas de Educación, Vol. 2, pp. 49-89. Montevideo: Universidad Católica del Uruguay.

Ravela, P., Leymonié. J, Viñas, J., Haretche, C. (2014). *La evaluación en las aulas de secundaria básica en cuatro países de América Latina.* En: Propuesta Educativa,

Año 23 Nº 41, pp. 20-45, Junio 2014. Buenos Aires: Facultad Latinoamericana de Ciencias Sociales.

Shepard, L. (2006). *La evaluación en el aula*. México: Instituto Nacional para la Evaluación de la Educación (INEE), 2008. Traducción del original en inglés, *Classroom Assesment;* en Brennan, R. (ed.) Educational Measurement, 4ta. Edición. ACE/Praeger Westport, 2006, pp. 623-46.

Wiggins, G. (1998). *Educative Assessment. Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.

Wiggins, G. y Mc Tighe, J. (2005). *Understanding by Design*. Segunda edición expandida. Alexandria, Virginia: ASCD.

Wiliam, D. (2011). *Embedded Formative Assessment*. Bloomington: Solution Tree Prees.

EPÍLOGO

La necesidad de políticas educativas que construyan un nuevo modelo para el trabajo docente

Llegamos al final del libro. Hemos realizado un largo recorrido por una amplia variedad de aspectos de la evaluación de aprendizajes en el aula. Como planteamos al comienzo, nuestra mayor aspiración al escribirlo ha sido lograr que algunos de los lectores docentes lo utilicen para trabajar con otros colegas en la revisión de sus prácticas de evaluación de enseñanza, y en la construcción de nuevas aproximaciones y experiencias.

Si bien el libro fue escrito para promover el intercambio y el desarrollo de prácticas innovadoras por parte de los lectores docentes, no se nos escapa, tal como lo planteamos en el primer capítulo, que hacer esto posible depende en buena medida de políticas orientadas al desarrollo profesional docente y de la construcción de una nueva concepción del trabajo en la profesión. Depende también, tal como fue planteado en el capítulo cinco, del desarrollo de una nueva matriz organizacional para el sistema educativo, que supere muchas de las características heredadas del modelo escolar creado en el marco de la sociedad industrial.

Lo anterior implica que, si bien los docentes tenemos cierto espacio de acción para introducir mejoras en nuestras formas de enseñar y de evaluar, no debemos caer en un voluntarismo que nos culpabiliza. Una educación más adecuada a las necesidades de los niños, niñas y jóvenes, así como a los requerimientos de la sociedad y la cultura en las que vivimos, depende de políticas educativas que promuevan una nueva forma de concebir, organizar y llevar adelante el trabajo docente. Esto es responsabilidad del sistema político, de las autoridades educativas y de las propias organizaciones de docentes. Es responsabilidad también de los directivos de las instituciones educativas públicas y privadas y, por supuesto, de nosotros mismos.

Una nueva concepción del trabajo docente es una tarea de fundamental importancia. Tanto entre muchos directivos de centros de enseñanza, como a nivel de las autoridades del sistema, se suele concebir la formación en servicio

como actividades y talleres que se realizan en forma aislada, normalmente en períodos de vacaciones de los estudiantes. No se alcanza a comprender la necesidad de un espacio permanente para la revisión, el intercambio y el desarrollo de innovaciones, que forme parte del trabajo cotidiano.

Tal como indicamos al comienzo del libro, tres son las características principales de un nuevo modelo para el trabajo en la profesión docente, que requieren voluntad política, acuerdos con los profesionales de la docencia, inversión de recursos y un cuidadoso diseño para su implementación:

1. el reconocimiento en el contrato y en la remuneración del tiempo de trabajo que necesariamente debe realizarse fuera del aula (y que es tan importante como el que se realiza dentro de ella);
2. la existencia de instancias institucionales frecuentes y regulares para el intercambio, la revisión de prácticas y el trabajo conjunto con otros colegas, enfocados en el desarrollo y experimentación de nuevas formas de trabajar en el aula (dejando atrás la concepción de la enseñanza como una actividad solitaria y escasamente observada y analizada);
3. la creación de nuevos roles de acompañamiento y orientación didáctica, radicados en los centros educativos (o para un conjunto de ellos, cuando son de pequeña escala), enfocados en el apoyo a la enseñanza, su preparación y evaluación, desempeñados por docentes experimentados y altamente competentes (lo cual implica superar los viejos dispositivos de inspección y supervisión orientados al control administrativo).

Asumiendo el riesgo de equivocarnos por falta de conocimiento, y sin desconocer la realización de experiencias específicas en pequeña escala, creemos no ha habido en la región en las últimas décadas una política de esta naturaleza. Ha habido esfuerzos variados por implementar programas de actualización y formación en servicio, otros dirigidos a promover experiencias innovadoras en los centros de enseñanza, algunas mejoras salariales, se han creado en algunos países espacios de "coordinación" o de reunión periódica de los docentes, entre otros. Pero ninguna política se ha planteado como objetivo central transformar en forma permanente el corazón de la concepción y organización del trabajo docente.

Simultáneamente, muchas otras políticas que han estado dirigidas a introducir nuevos instrumentos o apoyos para la docencia (como libros de texto, guías de enseñanza o nuevas tecnologías para el acceso a la información), han tenido un impacto muy limitado en las prácticas de enseñanza dentro de las aulas. Tal vez el Plan CEIBAL en Uruguay sea un ejemplo paradigmático

de ello. A lo largo de la última década se ha logrado que cada estudiante de educación primaria y media básica (1º a 9º grados) tenga un computador personal portátil (laptop), de su propiedad y entregado por el Estado.

Sin lugar a dudas este programa ha tenido un gran impacto en términos de inclusión digital y de facilitar el acceso de todos los niños y niñas (y sus familias) a nuevas fuentes de información y comunicación. Sin embargo, su impacto en la transformación de la forma de gestionar las clases ha sido bastante limitada.

Mientras que algunos docentes han desarrollado nuevas y efectivas formas de trabajo con las nuevas herramientas tecnológicas, otros simplemente no las utilizan en el aula, en tanto que muchos las incorporan a sus viejas prácticas de enseñanza. Presentamos un par de ejemplos, tomados de un estudio de observación de clases realizado algunos años atrás[1]. Una observación llamativa fue que muchas maestras iban guiando paso a paso lo que los alumnos debían hacer en su computadora, evitando que trabajasen a ritmos diferentes. Todos los alumnos entraban en la misma "página" al mismo tiempo. La maestra los hacía esperar hasta que todos estuviesen listos, y todos presionaban la tecla "ENTER" al mismo tiempo. Otro ejemplo llamativo fue la constatación de que, luego de buscar información en internet y tener un intercambio sobre la misma, los alumnos registraban sus notas en el cuaderno, copiando del pizarrón las conclusiones escritas por la maestra, en lugar de utilizar un procesador de textos.

Este tipo de hechos no debe sorprender ni implica una descalificación al docente. Se trata simplemente de una constatación de lo planteado en el capítulo 1: la enseñanza es una práctica cultural y modificarla requiere mucho tiempo y trabajo colaborativo. Los autores del trabajo analizado en ese capítulo[2] ponen ejemplos similares de situaciones en que una supuesta innovación es "absorbida" por la cultura existente. Relatan el caso de un profesor de matemática que, luego de ver cómo trabajaban los profesores de Japón, intentó utilizar el mismo enfoque, comenzando su clase proponiendo a los estudiantes un problema complejo cuya solución debían descubrir por sí mismos. Los estudiantes, que no habían visto el video, actuaron de acuerdo a su rol tradicional y se quedaron esperando a que el docente les explicara cómo resolver el problema.

---

1. Pérez Gomar, G. y Ravela, P., 2012. Impactos del Plan Ceibal en las prácticas de enseñanza en las aulas de primaria. Montevideo: Universidad Católica del Uruguay. Disponible en: http://www.ibertic.org/evaluacion/biblioteca/35/impactos-del-plan-ceibal-en-las-prácticas-de-enseñanza-en-las-aulas-de-primaria. Acceso: 27/01/17.
2. Hiebert y Stigler (2009: 99), ver referencia bibliográfica en el capítulo 1.

Esto mismo ocurrirá con la mayoría de los intentos de introducir nuevos elementos en la educación, mientras no vayan acompañados por una política dirigida a generar las condiciones para crear un ambiente y una cultura de trabajo innovadores en la profesión docente.

Los sistemas de enseñanza, como otros sistemas complejos, están compuestos por muchos elementos que interactúan y se refuerzan mutuamente. El todo es más que la suma de las partes. Por tanto, será difícil, sino imposible, modificar la enseñanza introduciendo nuevos textos, guías o elementos tecnológicos. Lo normal será que los docentes, en el mejor de los casos (salvo excepciones) adoptarán los nuevos elementos de un modo tal que encajen en el sistema de prácticas preexistente, en lugar de modificar dichas prácticas[3].

Cambiar la enseñanza es un cambio cultural. Requiere principalmente de un cambio en las concepciones y condiciones básicas del trabajo docente, de la mano con otros cambios en la matriz educativa, según fue planteado en el capítulo 5. El objetivo central debe ser hacer de la docencia una profesión en la que se experimenta, se crea y se innova en forma permanente. Para ello es necesario superar el aislamiento del trabajo docente y naturalizar la instancias de observación de clase y de revisión y preparación conjunta de materiales, así como el intercambio entre centros de enseñanza y al interior de los mismos (es notable lo poco que los docentes de un mismo establecimiento conocen acerca de lo que hacen sus colegas, tanto en el sector público como en el privado).

Para comprender mejor esta propuesta es útil reseñar rápidamente la forma en que en Japón[4] los docentes tienen naturalizado el trabajo colectivo para la preparación de clases y las visitas de observación mutuas. Todas las semanas, todos los docentes de educación primaria y media del país tienen un espacio de reunión (denominado "**kounaikenshuu**"), destinado a la preparación detallada de una clase sobre un tema para un curso específico. Por ejemplo, cómo trabajar triángulos equiláteros con niños de tercer grado de la escuela primaria. Los grupos están integrados por entre 5 y 7 docentes y las reuniones tienen una duración de entre tres y cuatro horas. Le dedican un semestre completo a diseñar una lección. El proceso de trabajo es más o menos como sigue:

a. comienzan por definir qué quieren que los estudiantes logren;
b. luego planifican detalladamente la lección, incluyendo aspectos tales como cuál será el problema disparador y cómo redactarlo, qué

3. Hiebert y Stigler (2009: 97), ver referencia bibliográfica en el capítulo 1.
4. Hiebert y Stigler (2009), capítulo 7, ver referencia bibliográfica en el capítulo 1.

materiales entregar a los estudiantes, qué preguntas hacer para promover la reflexión durante la clase, cómo organizar el uso del espacio en el pizarrón y cómo distribuir el tiempo, qué caminos de resolución probablemente ensayarán los estudiantes y qué dificultades encontrarán, y, finalmente, cómo cerrar la lección;

c. en tercer lugar algunos de los docentes del grupo enseñan esa lección en sus aulas, en tanto que los demás concurren como observadores;
d. a partir de la experiencia implementada y observada, se analiza lo ocurrido en las aulas y se evalúa la lección, se la revisa y se la rediseña;
e. luego se vuelve a enseñar la lección revisada, por los mismos u otros docentes, también con observación de los demás colegas;
f. finalmente, se evalúa lo ocurrido, se elabora la versión final y se comparte el resultado con los colegas del resto del sistema educativo.

Lo que hace especialmente interesante este sistema es que cada grupo de docentes trabaja sobre un tema distinto y que, al final del año, todas las lecciones elaboradas son publicadas y puestas a disposición de todos los docentes del país. De esta manera, se construye un formidable sistema de generación y sistematización de conocimiento práctico, desde la propia profesión. No se trata de esfuerzos o iniciativas aisladas, sino de un emprendimiento colaborativo del conjunto de los docentes, sostenido desde la política educativa y desde la estructura del propio sistema, que hace posible tanto las instancias de reunión, como los tiempos para la observación de clases entre pares y la divulgación de las producciones de los grupos.

Un política de este estilo es lo que, creemos, estamos necesitando ver en nuestra región.

Este libro ha sido escrito en Uruguay, pero pensando en los docentes de toda América Latina, maestros y profesores de educación primaria y media. Nuestra intención ha sido construir una herramienta de trabajo que ayude a los colegas docentes a reflexionar, revisar y mejorar sus formas de evaluar a los estudiantes. Pretende ser un aporte para el trabajo en el aula en los temas de evaluación. También ha sido pensado como una herramienta para quienes tienen a su cargo la formación de docentes, tanto la inicial como la formación en servicio.

Modificar nuestra forma de "hacer las cosas" en el aula requiere una cierta dosis de humildad. El libro invita al lector a revisar sus prácticas de enseñanza, a partir de una mirada reflexiva sobre sus propias evaluaciones. Nos proponemos abrir ventanas hacia aspectos de la evaluación que, a partir de nuestra experiencia, de nuestras investigaciones y de nuestro intercambio con colegas docentes, creemos que podrían hacerse mejor. Cada lector, a partir de su trayectoria, formación y experiencia acumulada, encontrará distintos aspectos para revisar y mejorar.

El libro no es un manual de evaluación, ni un libro de teoría de la evaluación, ni tampoco un informe de investigación, pero tiene algo de cada uno de ellos. A lo largo del mismo el lector encontrará reflexiones y análisis conceptuales sobre la evaluación; resultados de trabajos de investigación en torno a la evaluación en el aula en la región; ejemplos de trabajos y experiencias innovadoras en evaluación realizadas en distintos países.

A lo largo del libro se abordan los siguientes temas principales:

- una aproximación conceptual a la evaluación en el aula y a las dificultades para modificar las prácticas de enseñanza;
- herramientas para analizar los componentes centrales de una actividad de evaluación y los procesos de pensamiento que están involucrados en su resolución;
- sugerencias y ejemplos para transformar las pruebas que utilizamos habitualmente, en situaciones más ricas, motivadoras y desafiantes para los estudiantes, a través de propuestas de evaluación auténtica;
- pistas de trabajo para incorporar la evaluación formativa al trabajo cotidiano, analizando las maneras de formular preguntas en clase y los modos de corregir y devolver trabajos y pruebas. Se enfatiza el concepto de "devolución" y se brindan sugerencias para promover de manera efectiva instancias de auto-evaluación y co-evaluación de los estudiantes;
- un análisis de los reglamentos de evaluación vigentes en varios países de la región y del modo en que estos determinan prácticas de calificación cuya consecuencia principal es generar una cultura en la que los estudiantes están más orientados a las calificaciones que al aprendizaje. Se plantea la necesidad de pensar en un modelo alternativo y un conjunto de sugerencias para mejorar las formas en que habitualmente calificamos a los estudiantes.

ISBN: 978-607-8558-39-1











